Ocean data assimilation

OACOS master's program
February 7th, 2017

Emmanuel COSME
Université Grenoble Alpes
IGE
Grenoble

Acknowledgements

* This presentation has been set up thanks to the particular contributions of Pierre Brasseur, Charles-Emmanuel Testut, Eric Blayo, Jean-Michel Brankart, Pierre Antoine Bouttier, Clément Ubelmann.
* These names hide many others who indirectly contributed, particularly from LGGE/MEOM, Mercator-Océan, LJK/MOISE, and NASA/JPL.
* And thank you for the invitation to give this course.

Scope of this lecture

* This DA lecture mostly deals with:
  * the ocean circulation
  * the ocean primary production
* This lecture does not address:
  * ocean wave forecasting
  * tidal/storm surge forecasting
  * ocean chemistry and water quality
  * Fish, whales, sharks, jellyfish…

Scope of this lecture

* This lecture is biased towards realistic applications:
  * Realistic models;
  * Real observations;
  * Practical implementation of DA;
  * And a very limited amount of theory.

Operational oceanography: the primary user of ocean data assimilation

* Operational oceanography started about 20 years ago;
* The main goal is real-time monitoring and prediction of the state of the ocean, including:
  * Currents (shipping, sea operations, regattas…)
  * Primary production (marine resources, fishing)
  * Sea ice (shipping)
  * Temperature (climate, weather forecasting…)
* Like weather forecast centers, OO centers turn to provide useful information to scientists: reanalyses, targeted forecasts for field campaigns, etc.

Mercator-Océan

* The French center of OO;
* Created in 1995;
* Located in the area of Toulouse, about 50 agents;
* officially appointed by the European Commission on 11 November, 2014 to implement and operate the Copernicus Marine Service (CMEMS).

Mercator-Océan and research groups

* To develop its operational system, Mercator-Océan relies on the research community in the labs. In France, these are primarily (non-exhaustive list in almost arbitrary order):
  * IGE/MEOM (Grenoble)
  * LOCEAN (Paris)
  * LPO (Brest)
  * LEGOS (Toulouse)
  * CERFACS (Toulouse)
  * Météo-France (Toulouse)
  * etc

Web sites

* Mercator-Océan: http://www.mercator-ocean.fr/
* CMEMS : http://marine.copernicus.eu/
* GODAE Oceanview: https://www.godae-oceanview.org/
* DRAKKAR project: http://www.drakkar-ocean.eu/
* GFDL Ocean modeling: http://ocean-modeling.org/
* Coriolis data center: http://www.coriolis.eu.org/

Textbooks

* Advanced data assimilation for Geosciences, Eds. E. Blayo, M. Bocquet & E. Cosme, Oxford, 2014
* Data assimilation, Making sense of observations, Eds W. Lahoz, B. Khattatov & R. Ménard, Springer, 2010
* Ocean Weather Forecasting, Eds. E. Chassignet & J. Verron, Springer, 2006

Outline

* Ocean models
* Observations of the ocean
* Ocean DA using Ensemble Kalman filters
* Ocean DA using variational methods
* Future challenges

* Time allowing: Mercator-Océan operational DA system

Ocean models

* Primitive equations
* Scales
* Horizontal discretization
* Uncertainties
* Biogeochemistry

Primitive equations

$$
\begin{aligned}
\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y} + w\frac{\partial u}{\partial z} &= fv - \frac{1}{\rho}\frac{\partial p}{\partial x} + K_u\frac{\partial^2 u}{\partial z^2} \\
\frac{\partial v}{\partial t} + u\frac{\partial v}{\partial x} + v\frac{\partial v}{\partial y} + w\frac{\partial v}{\partial z} &= -fu - \frac{1}{\rho}\frac{\partial p}{\partial y} + K_v\frac{\partial^2 v}{\partial z^2} \\
-\frac{\partial p}{\partial z} &= \rho g
\end{aligned}
$$

$$\mathrm{div}\ \vec{u} = 0$$

$$\rho\frac{DS}{Dt} = \mathrm{div}\ (K_S \mathrm{grad}\ S)$$

$$\rho C_v\frac{DT}{Dt} = \mathrm{div}\ (K_T \mathrm{grad}\ T)$$

$$\rho = \rho(T, S, p)$$

+ auxiliary conditions

Conservation of:

- momentum
- Mass
- Salt
- Temperature

Equation of state

Primitive equations

Conservation of:

$$\frac{\partial u}{\partial t} + \boxed{u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y} + w\frac{\partial u}{\partial z}} = fv - \frac{1}{\rho}\frac{\partial p}{\partial x} + K_u\frac{\partial^2 u}{\partial z^2}$$

$$\frac{\partial v}{\partial t} + \boxed{u\frac{\partial v}{\partial x} + v\frac{\partial v}{\partial y} + w\frac{\partial v}{\partial z}} = -fu - \frac{1}{\rho}\frac{\partial p}{\partial y} + K_v\frac{\partial^2 v}{\partial z^2}$$

$$-\frac{\partial p}{\partial z} = \rho g$$

Nonlinear terms

- momentum

$$\mathrm{div}\ \vec{u} = 0$$

- Mass

$$\rho\frac{DS}{Dt} = \mathrm{div}\ (K_{\mathrm{S}}\mathrm{grad}\ S)$$

- Salt

$$\rho C_{\mathrm{v}}\frac{DT}{Dt} = \mathrm{div}\ (K_{\mathrm{T}}\mathrm{grad}\ T)$$

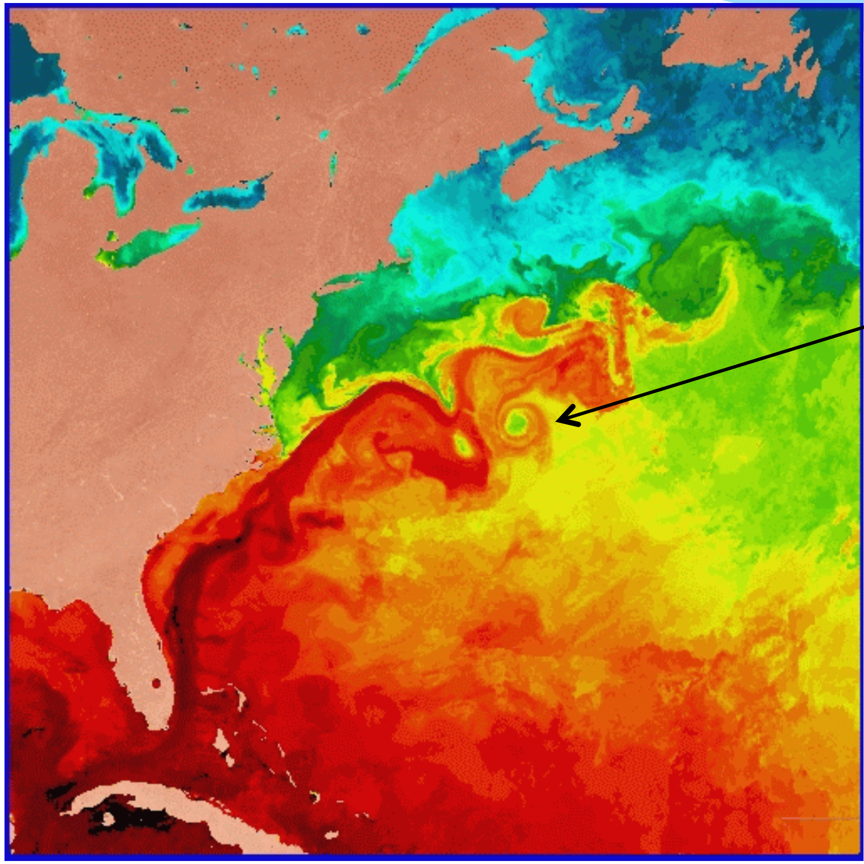
- Temperature

$$\rho = \rho(T, S, p)$$

Equation of state

+ auxiliary conditions

Primitive equations



Due to nonlinear terms

Primitive equations

* Why does this matter for DA?

  * Most tractable DA methods are designed for linear or weakly nonlinear systems;
  * All scales are involved and coupled in the dynamics. Representing the circulation accurately requires high-resolution (therefore expensive) models.

Scales

Phenomenon	Length scale L	Velocity scale U	Time scale T
Atmosphere:			
Sea breeze	5–50 km	1–10 m/s	12 h
Mountain waves	10–100 km	1–20 m/s	Days
Weather patterns	100–5000 km	1–50 m/s	Days to weeks
Prevailing winds	Global	5–50 m/s	Seasons to years
Climatic variations	Global	1–50 m/s	Decades and beyond
Ocean:			
Internal waves	1–20 km	0.05–0.5 m/s	Minutes to hours
Coastal upwelling	1–10 km	0.1–1 m/s	Several days
Large eddies, fronts	10–200 km	0.1–1 m/s	Days to weeks
Major currents	50–500 km	0.5–2 m/s	Weeks to seasons
Large-scale gyres	Basin scale	0.01–0.1 m/s	Decades and beyond

Scales

Scales particularly relevant for weather predictions and important for climate too.

Phenomenon	Length scale L	Velocity scale U	Time scale T
Atmosphere:			
Sea breeze	5–50 km	1–10 m/s	12 h
Mountain waves	10–100 km	1–20 m/s	Days
Weather patterns	100–5000 km	1–50 m/s	Days to weeks
Prevailing winds	Global	5–50 m/s	Seasons to years
Climatic variations	Global	1–50 m/s	Decades and beyond
Ocean:			
Internal waves	1–20 km	0.05–0.5 m/s	Minutes to hours
Coastal upwelling	1–10 km	0.1–1 m/s	Several days
Large eddies, fronts	10–200 km	0.1–1 m/s	Days to weeks
Major currents	50–500 km	0.5–2 m/s	Weeks to seasons
Large-scale gyres	Basin scale	0.01–0.1 m/s	Decades and beyond

Scales

* The scale of eddies is set by the Rossby radius of deformation:

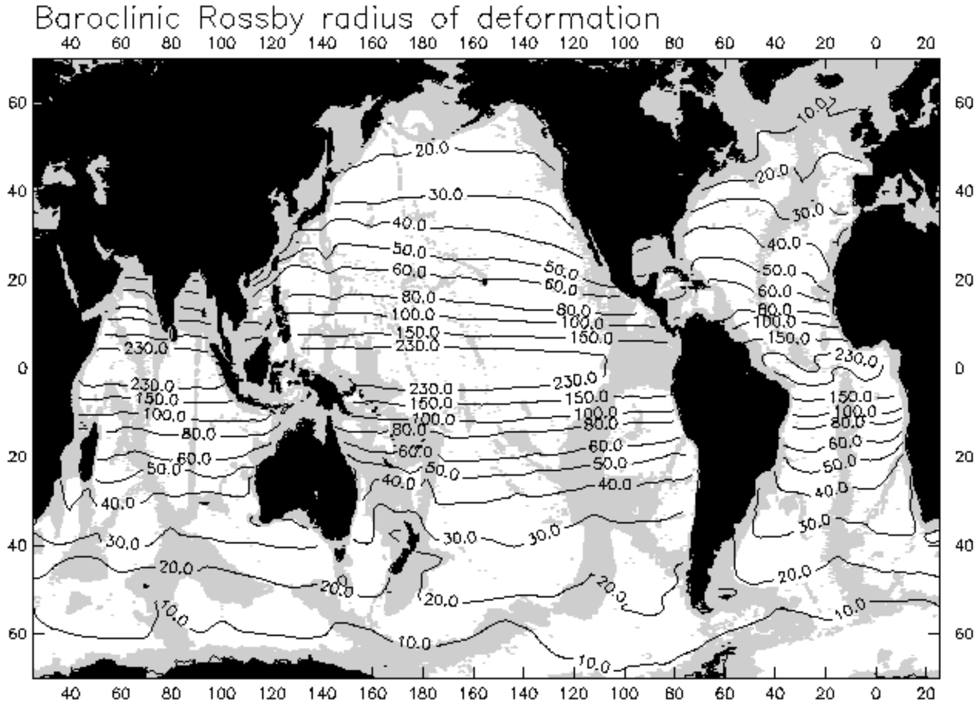
$$L_\rho = \frac{NH}{2\Omega}$$

N: Brunt-Vaïsala frequency

H: layer thickness

Ω: Earth rotation

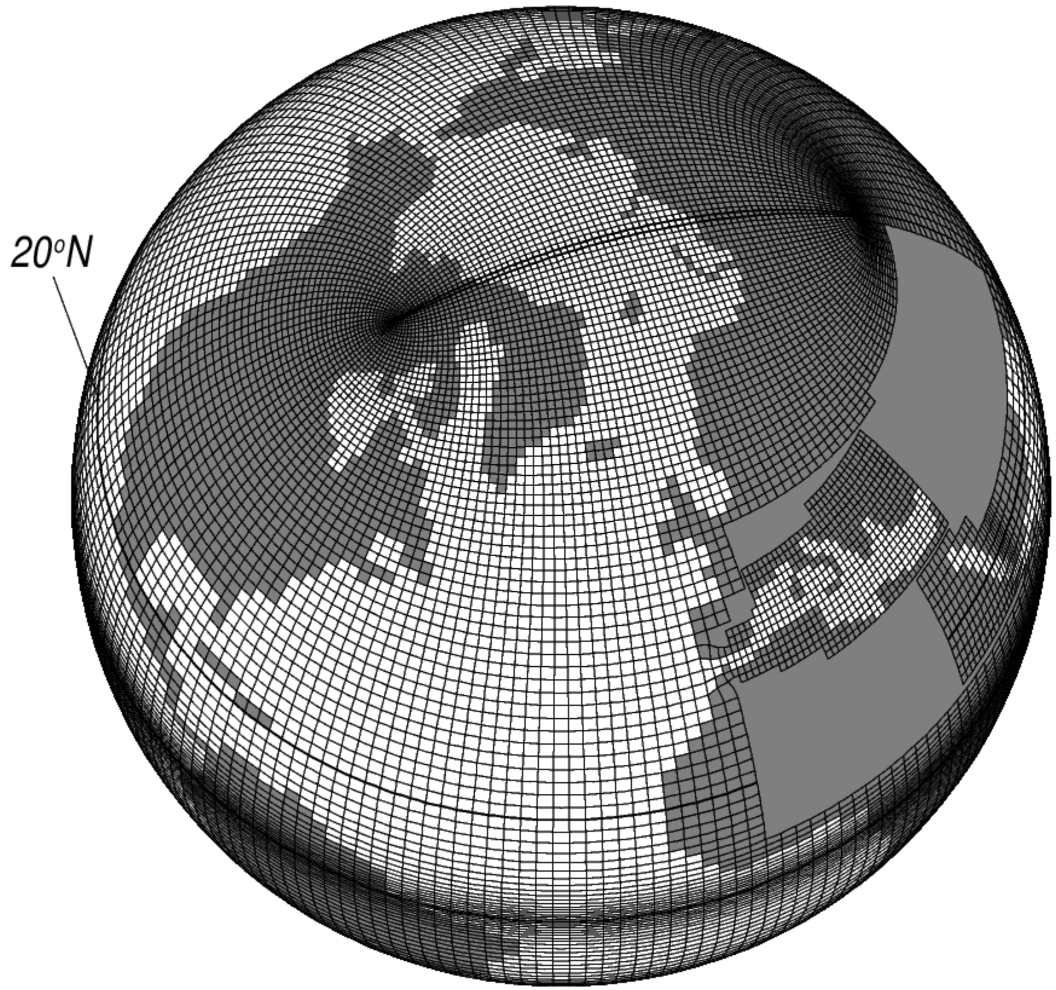
* ~30 km in the ocean, ~1000 km in the atmosphere
* Ocean weather simulations require high resolution models!



(Chelton et al, 1998)

Horizontal discretisation

* Figure: NEMO ORCA2 grid (2°)
* In 2015, operational version at 1/12° at Mercator-Océan
* Regional configuration at higher resolutions
* Resolution is pushed ahead…

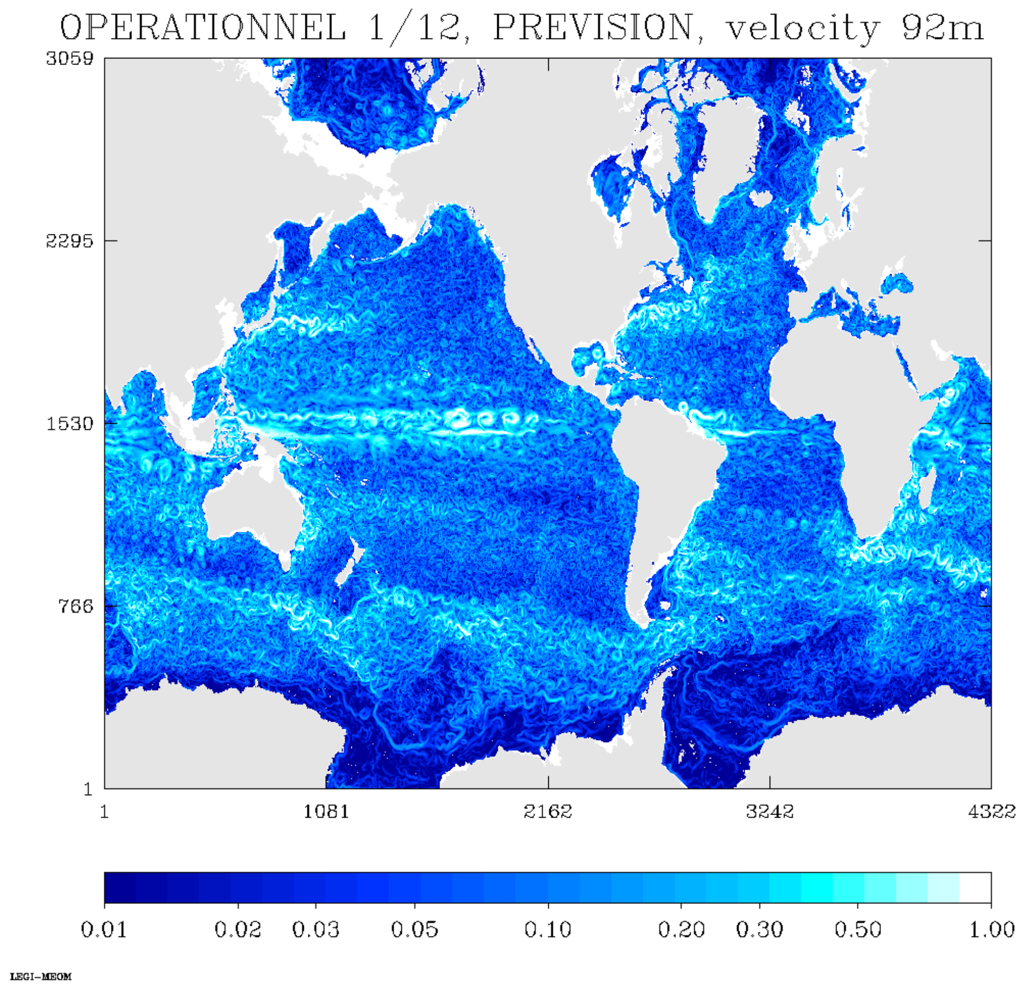


Horizontal discretisation

* Mercator operational model: NEMO 1/12°
* Number of gridpoints:

$$4322 \times 3059 \times 75 \sim 10^9$$

* 1 year of simulation costs 414 Gb memory, 90000 CPU hours, 1Tb storage (daily outputs)

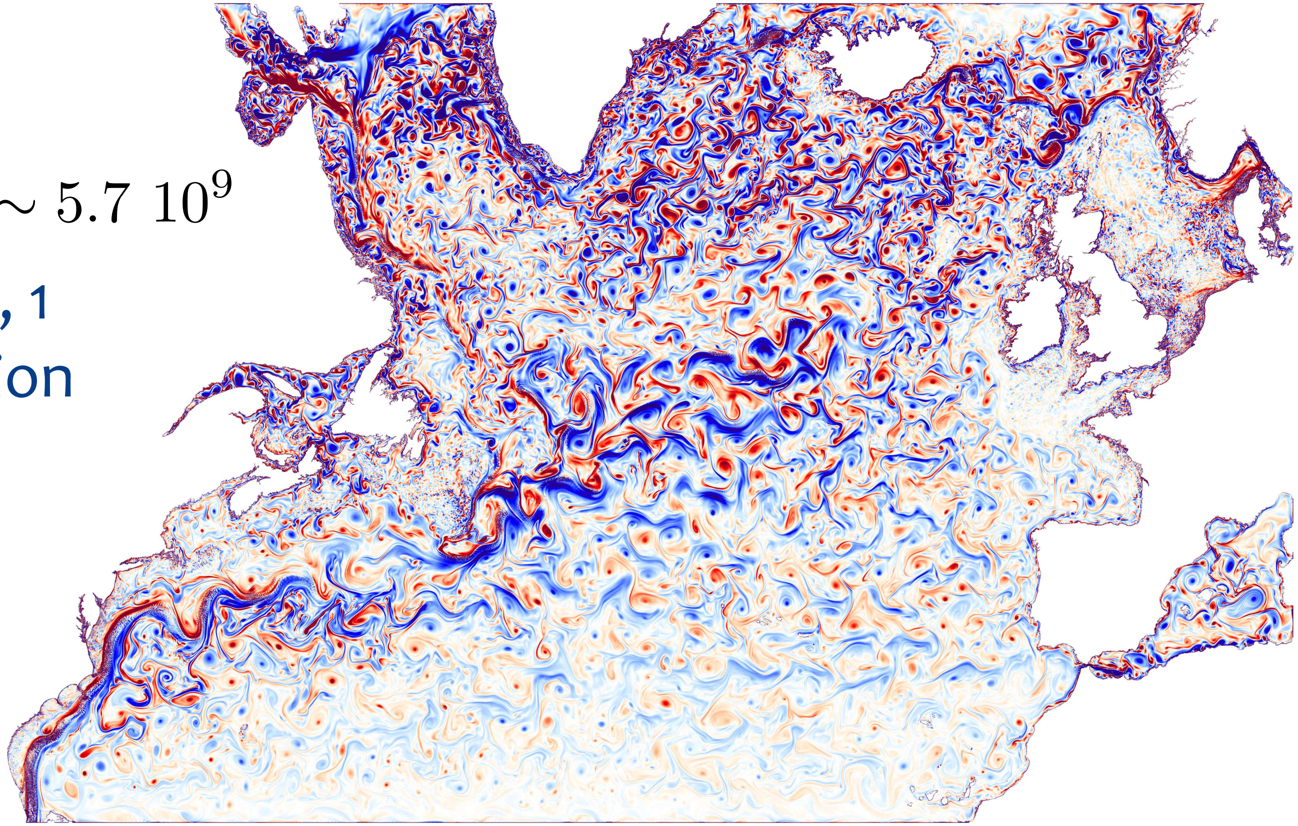


Horizontal discretisation

* NATL60
* Gridpoints:

$$5454 \times 3474 \times 300 \sim 5.7\ 10^9$$

* 13000 processors, 1 month of simulation takes 1 day



Horizontal discretisation

* Why does this matter for ocean DA?

  * The higher the resolution, the more expensive the model.
  * 4DVar needs iterations, EnKF requires an ensemble and accurate error covariances.
  * A huge volume of observations is needed to control such models.

Uncertainties

* Unresolved scales and parameterizations
* Forcings and boundary conditions

Uncertainties due to unresolved scales and parameterizations

* Example of a generally ignored effect: the state equation
  * Let <A> be the average value of A in the model gridpoint;
  * The model computes <T>, <S> from the conservation equations
  * Then computes density as:

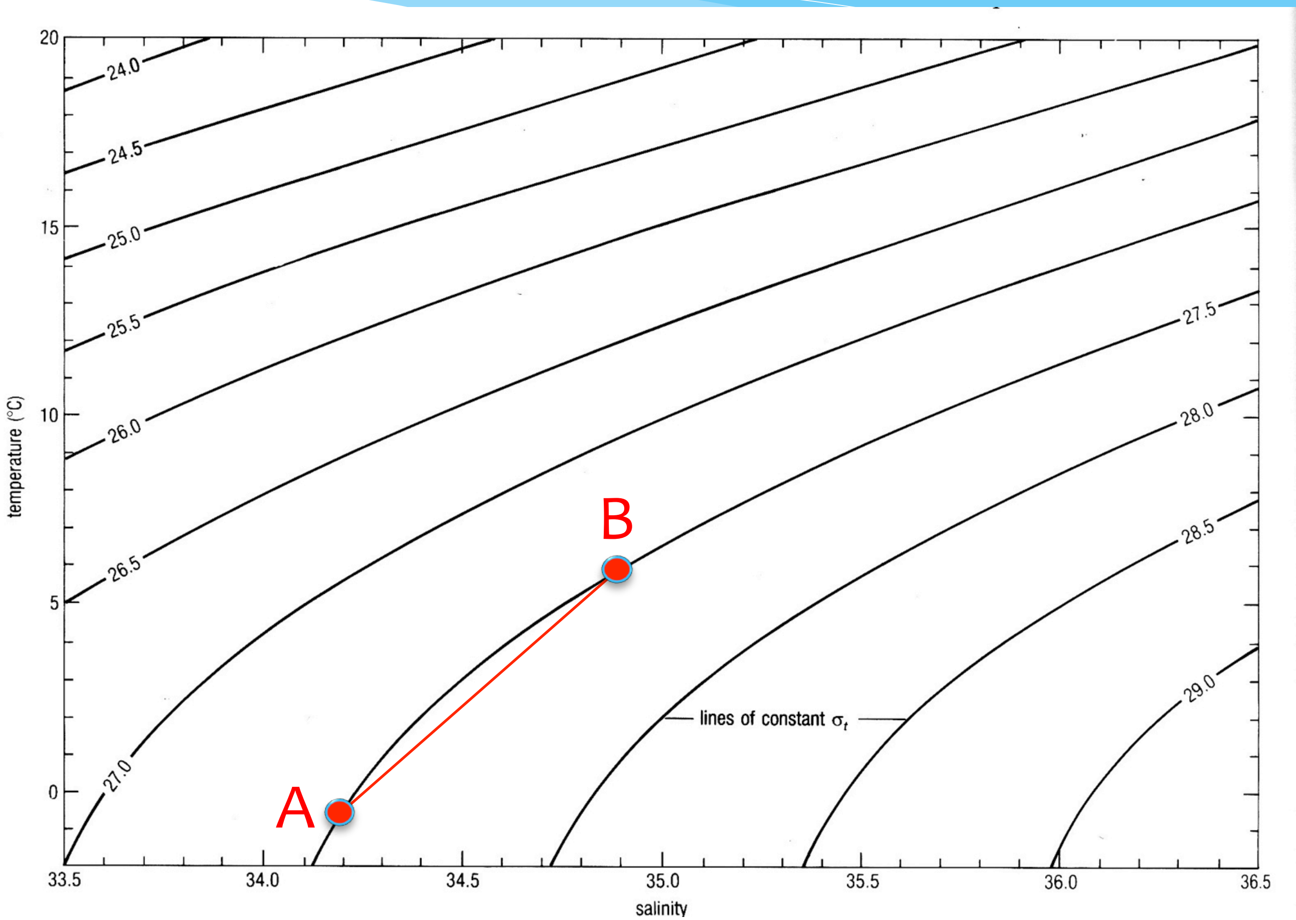
$$\rho = \rho(< T >, < S >)$$

* Which is different from:

$$\rho = < \rho(T, S) >$$

Ocean models

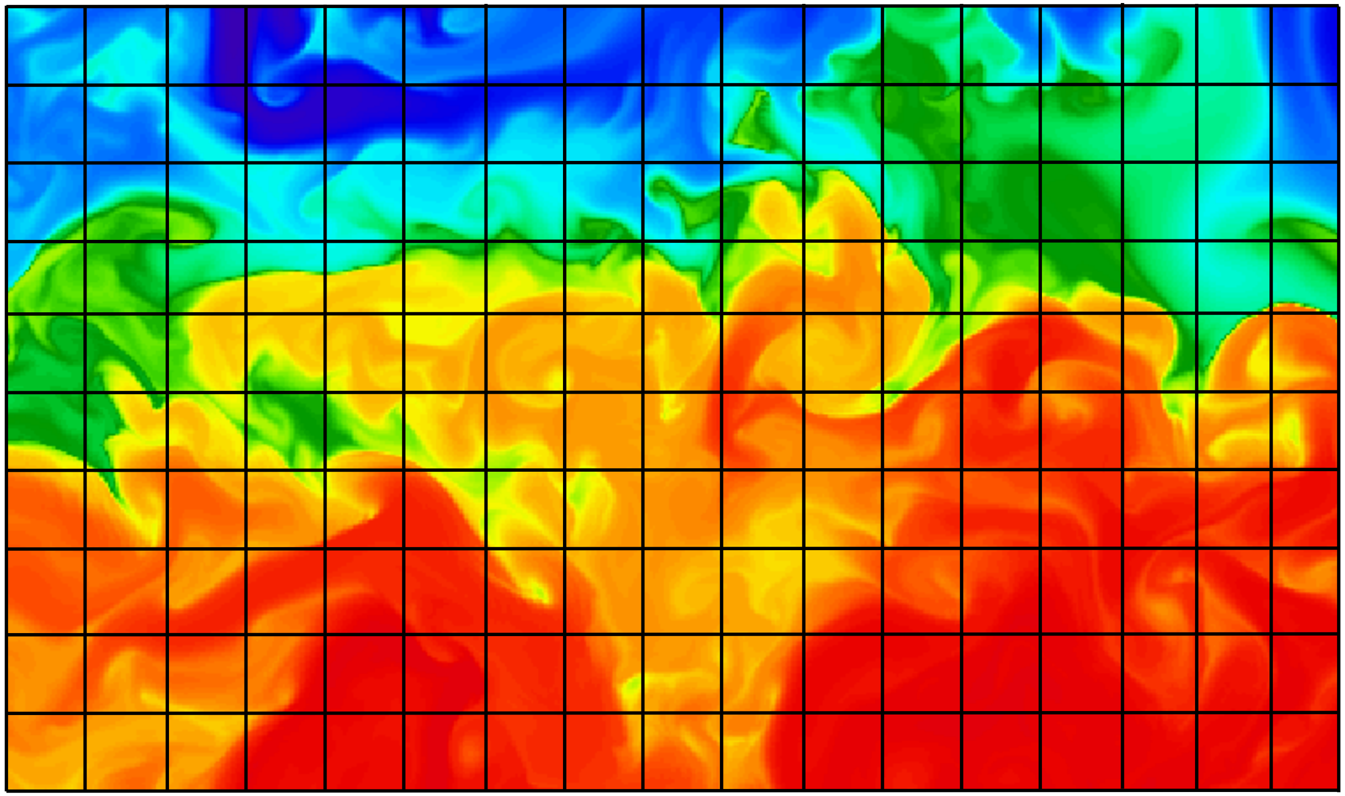
Uncertainties due to unresolved scales and parameterizations



$$\rho(A+B) \neq \rho(A) + \rho(B)$$

Uncertainties due to unresolved scales and parameterizations

A realistic temperature field and a possible model grid



Uncertainties due to unresolved scales and parameterizations

Idea:

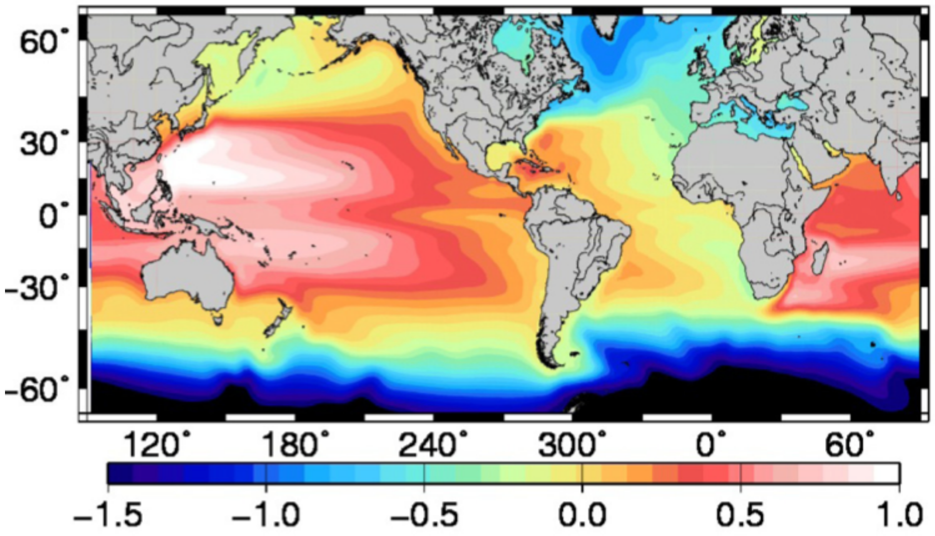
- represent the sub-grid variability in T and S with an ensemble, using stochastic (random) perturbations;
- Compute density for each (T, S) pair;
- Compute the density mean.

➡ Estimate $\rho = < \rho(T,S) >$ instead of $\rho = \rho(< T >, < S >)$

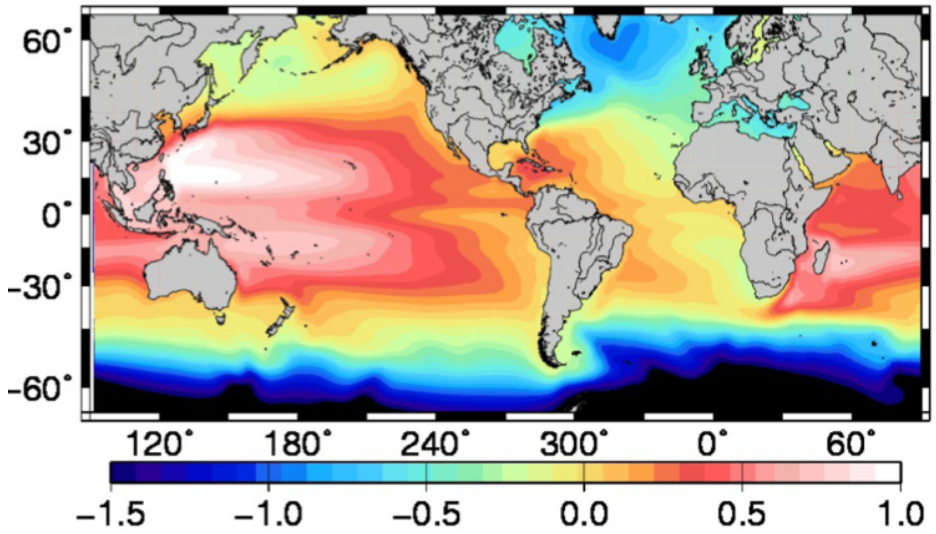
Uncertainties due to unresolved scales and parameterizations

Fields of SSH from NEMO, ORCA2 (gridmesh 2°)

$$\rho = \rho(< T >, < S >)$$



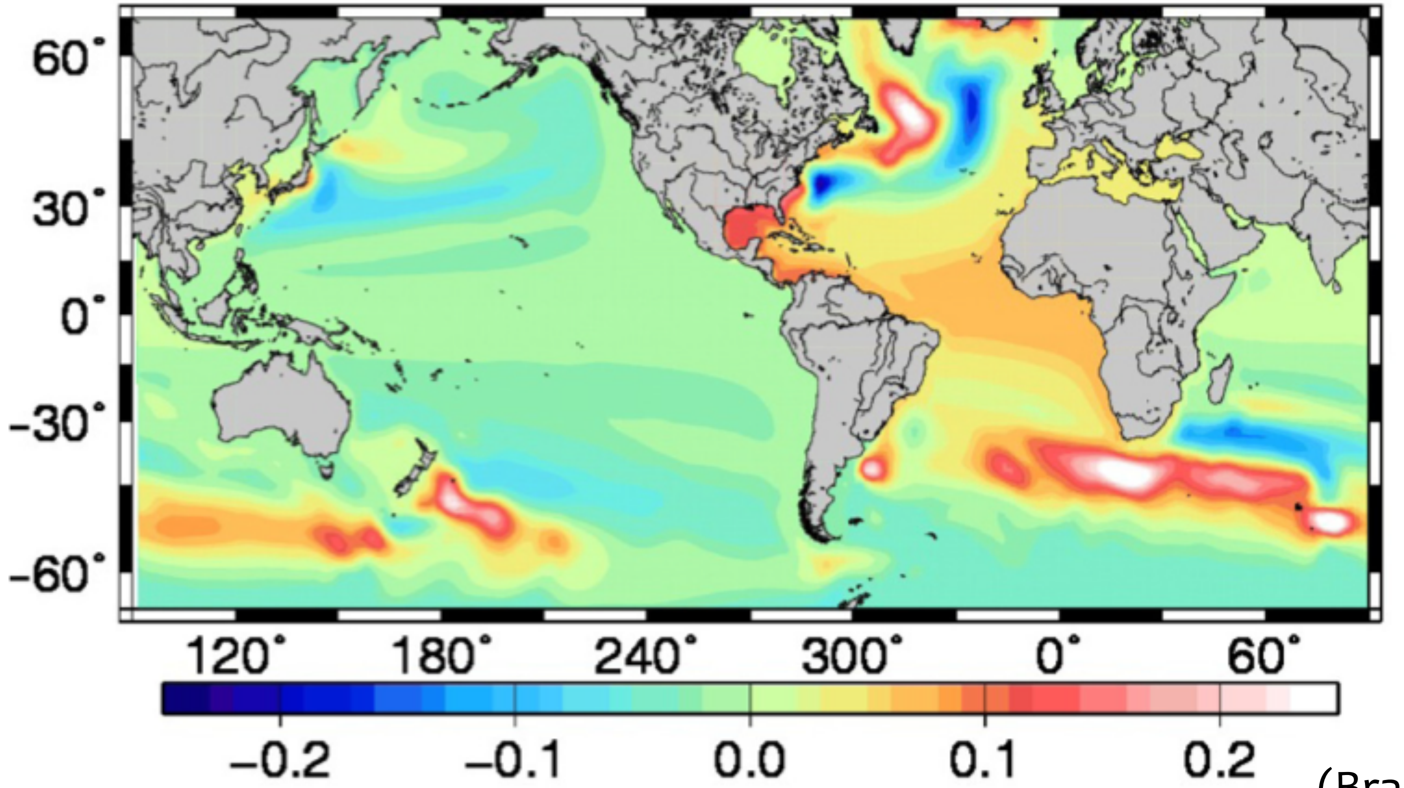
$$\rho =< \rho(T, S) >$$



(Brankart, 2013)

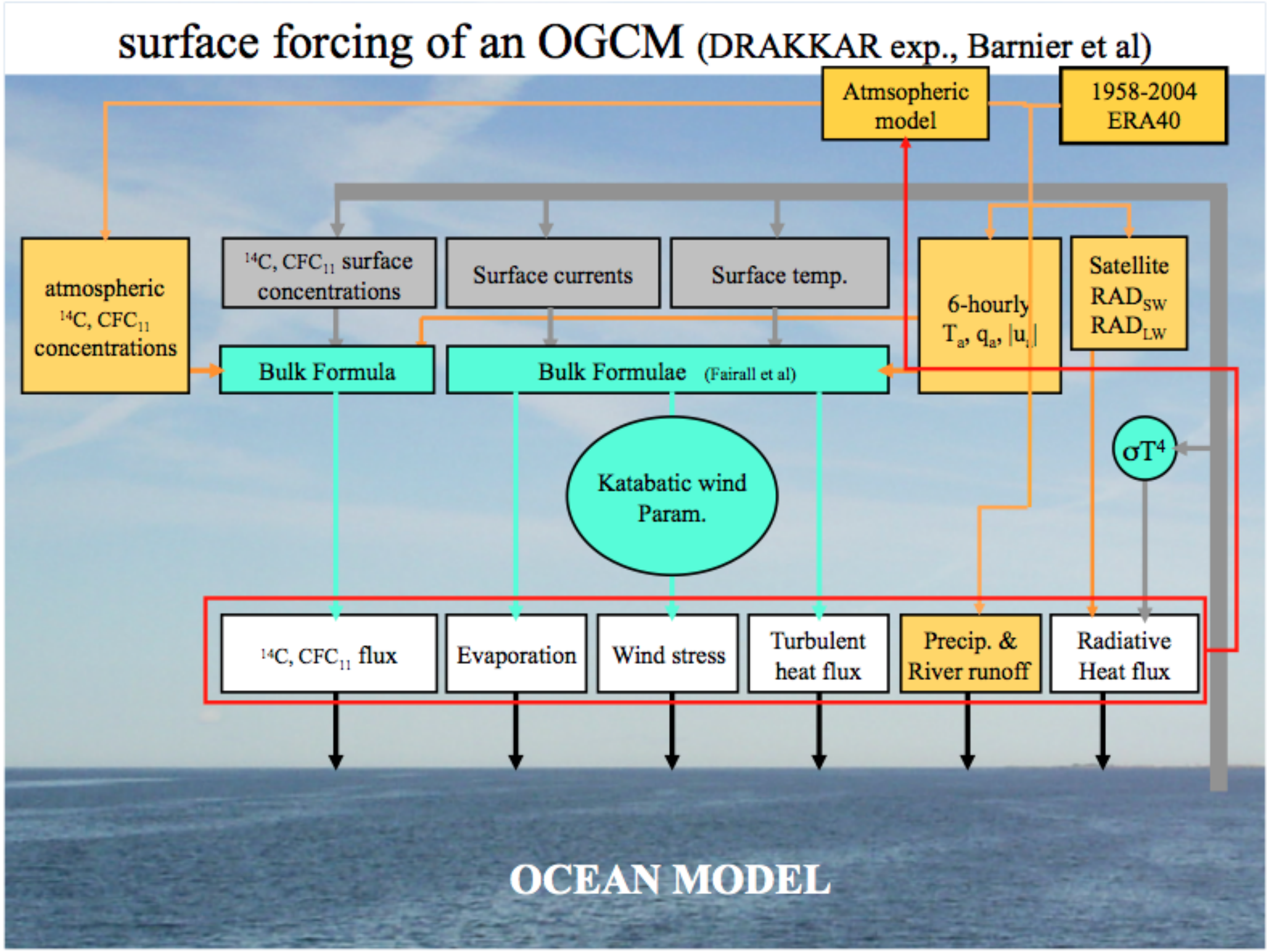
Uncertainties due to unresolved scales and parameterizations

Difference



(Brankart, 2013)

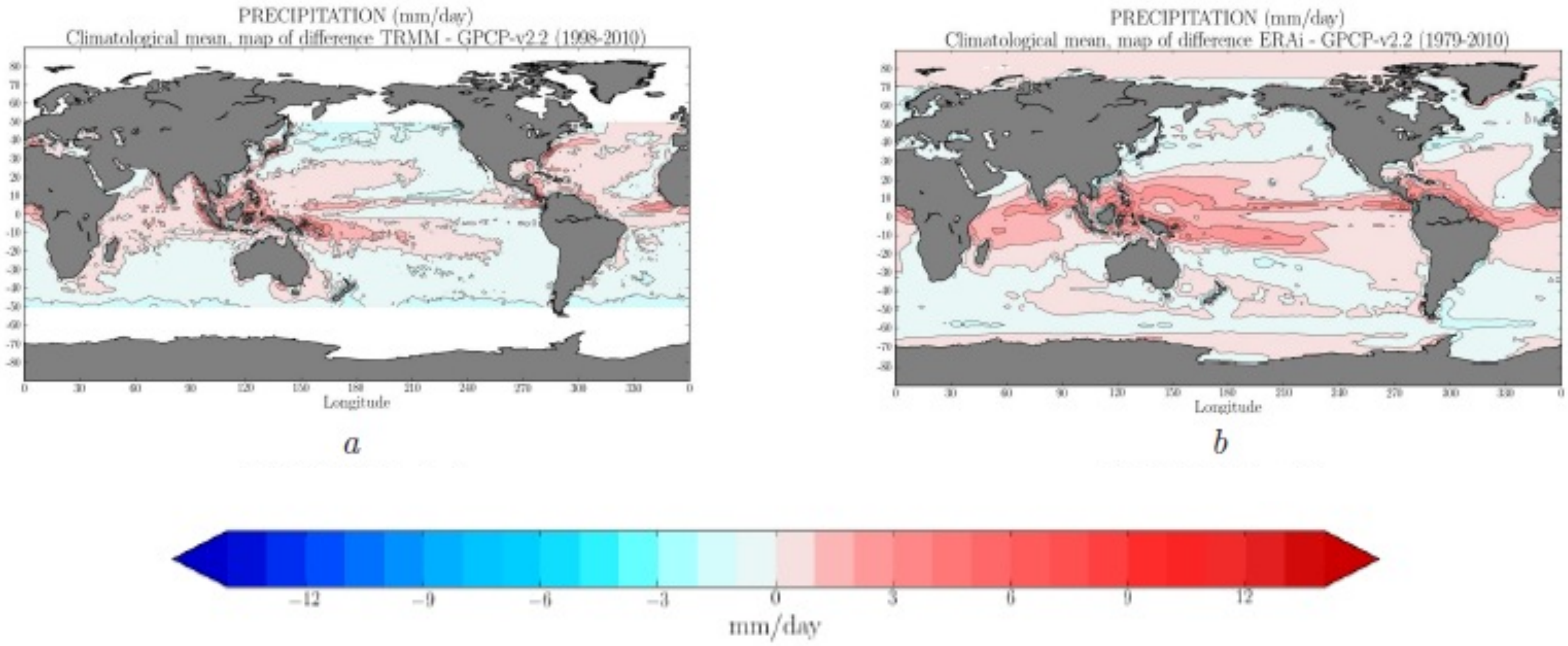
Uncertainties due to boundary conditions



Yellow: atmospheric
Grey: oceanic
Green: parameterizations
White: physical processes

Ocean models

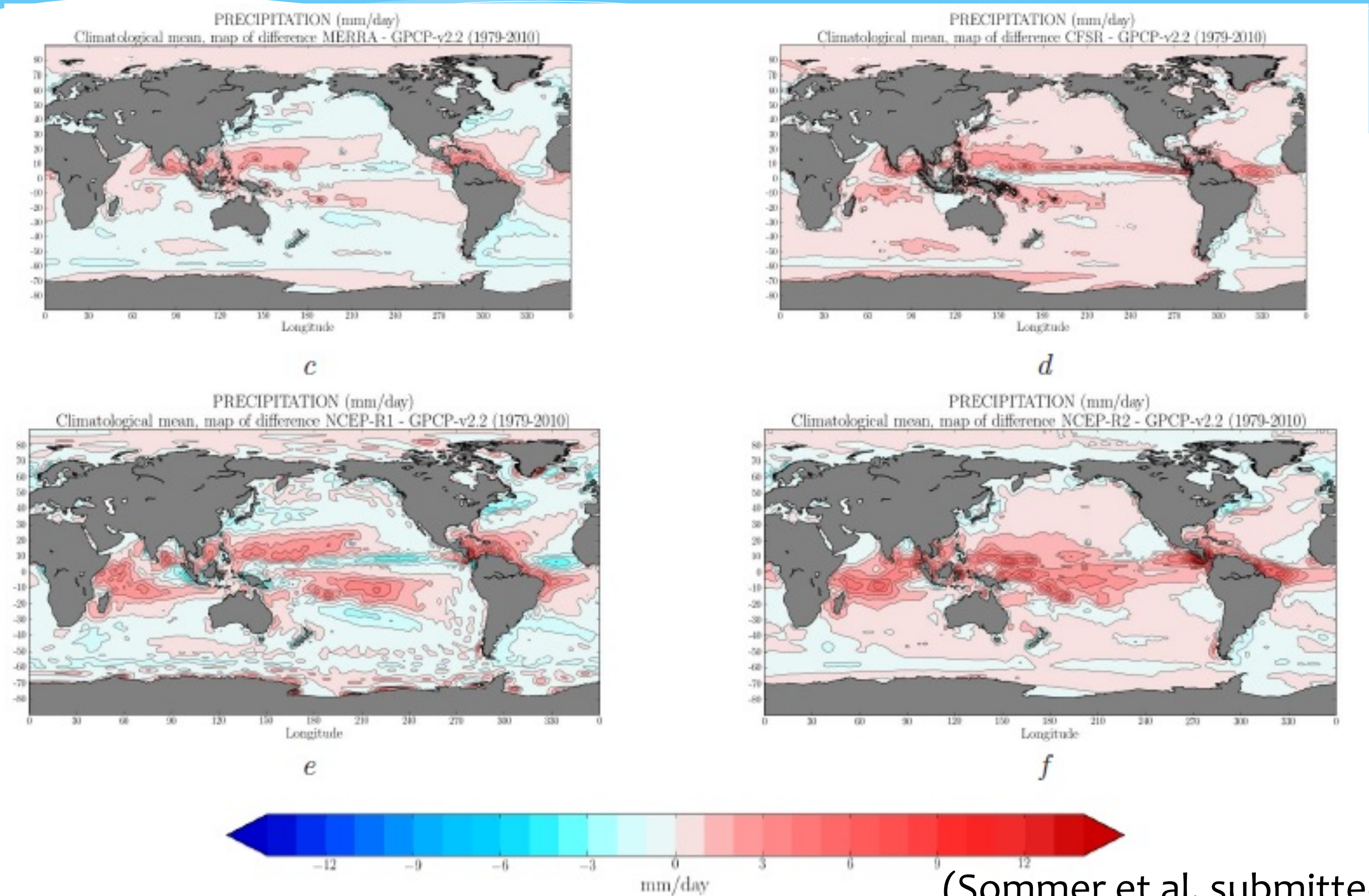
Uncertainties due to boundary conditions



(Sommer et al, submitted)

Ocean models

Uncertainties due to boundary conditions



(Sommer et al, submitted)

Ocean models: uncertainties

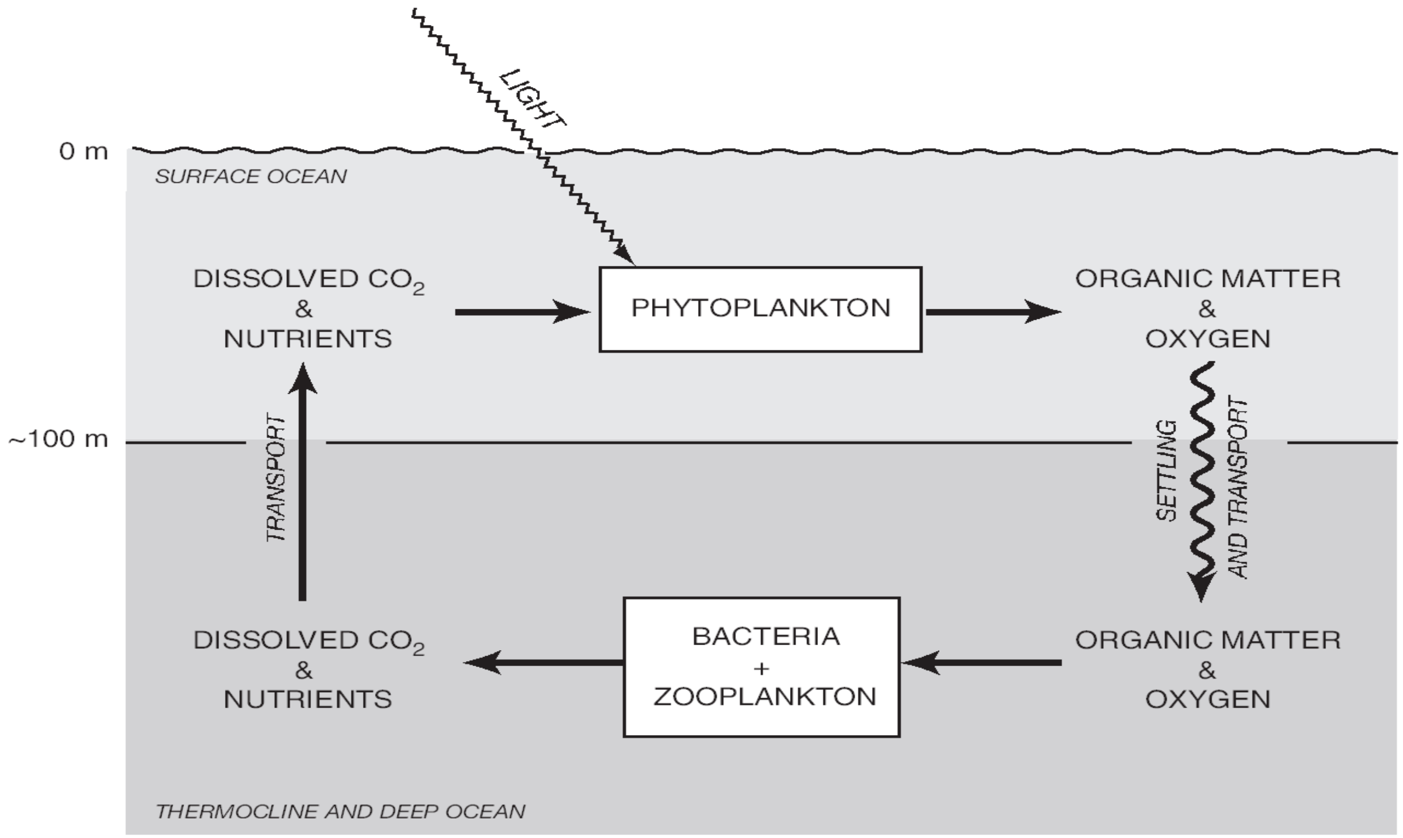
* Why does this matter for ocean DA?

  * Models has many sources of uncertainty;
  * To provide the best representation of the ocean state, models must be constrained by observations with DA;
  * To set up the DA system correctly, one must identify at best the various sources of errors and parameterize their impact;

  * DA can "guide" models, but also help in reducing the original uncertainties (e.g., by estimating parameters)

Ocean models

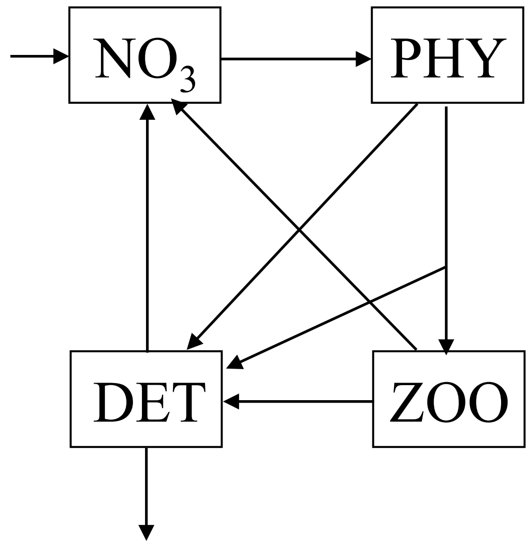
Biogeochemistry



Ocean primary production is a key piece of the ocean life and the carbon cycles.

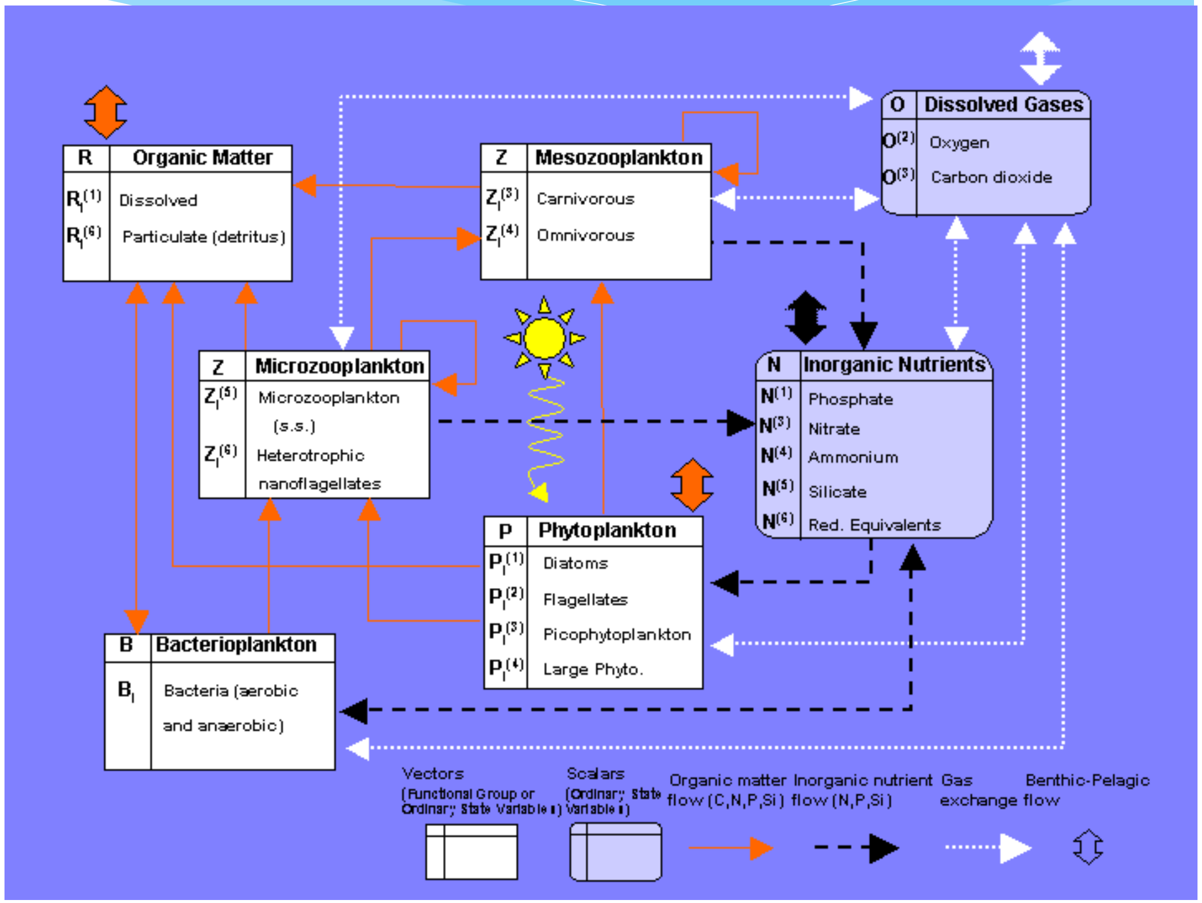
Biogeochemistry

* Simple NPZD ecosystem model (Nutrients, Phyto, Zoo, Detritus)
* 10-30 tunable parameters:
  * Growth rate, mortality
  * Sedimentation speed
  * Etc
* Already challenging for assimilation



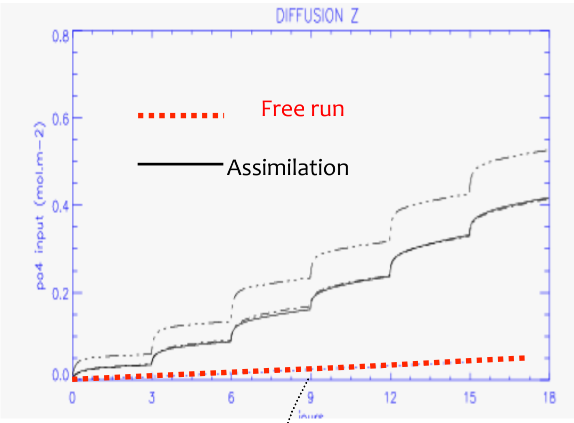
Biogeochemistry

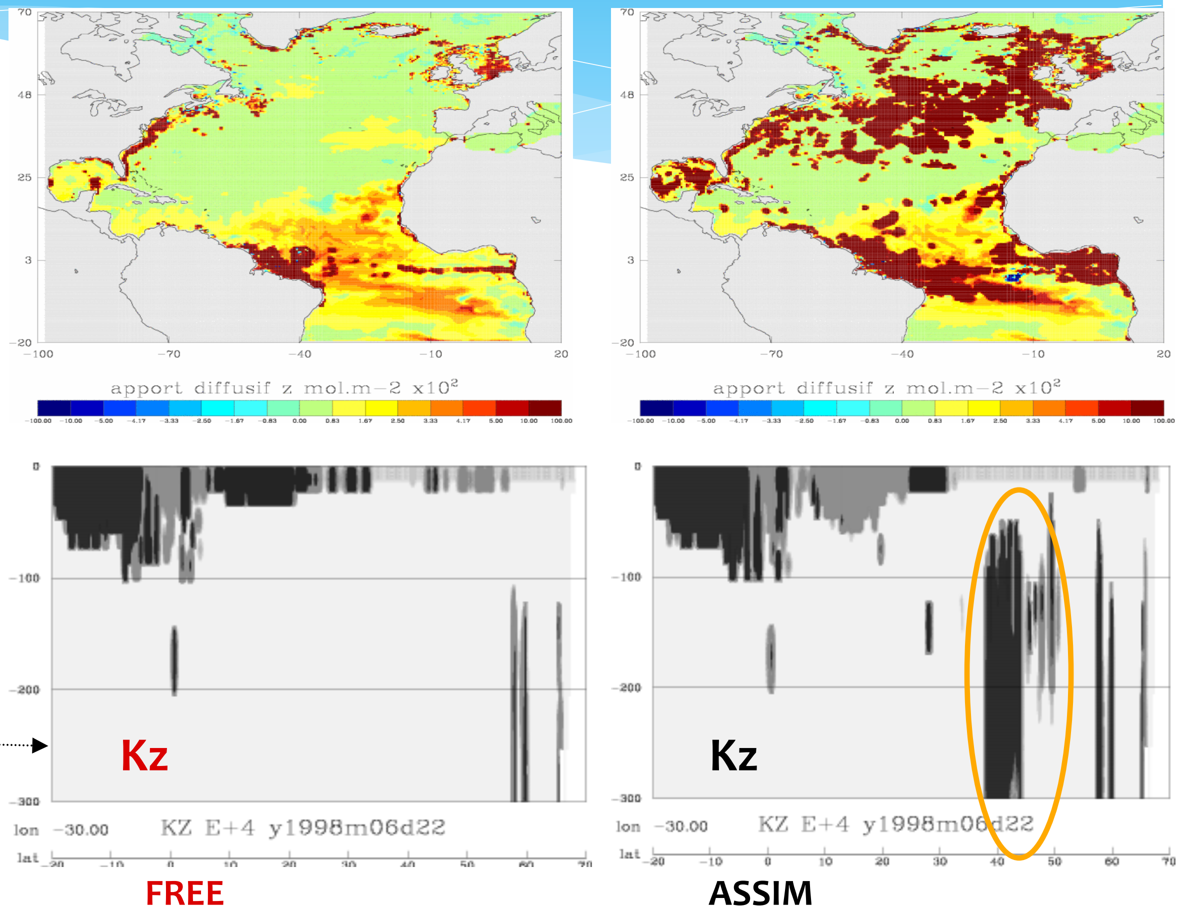
* Generic pelagic ecosystem models
* More than 100 tunable parameters



(Vichy et al, 2007)

Biogeochemistry





* An "bad" assimilation in the dynamics can be detrimental to biology

(Berline et al, 2005)

Biogeochemistry

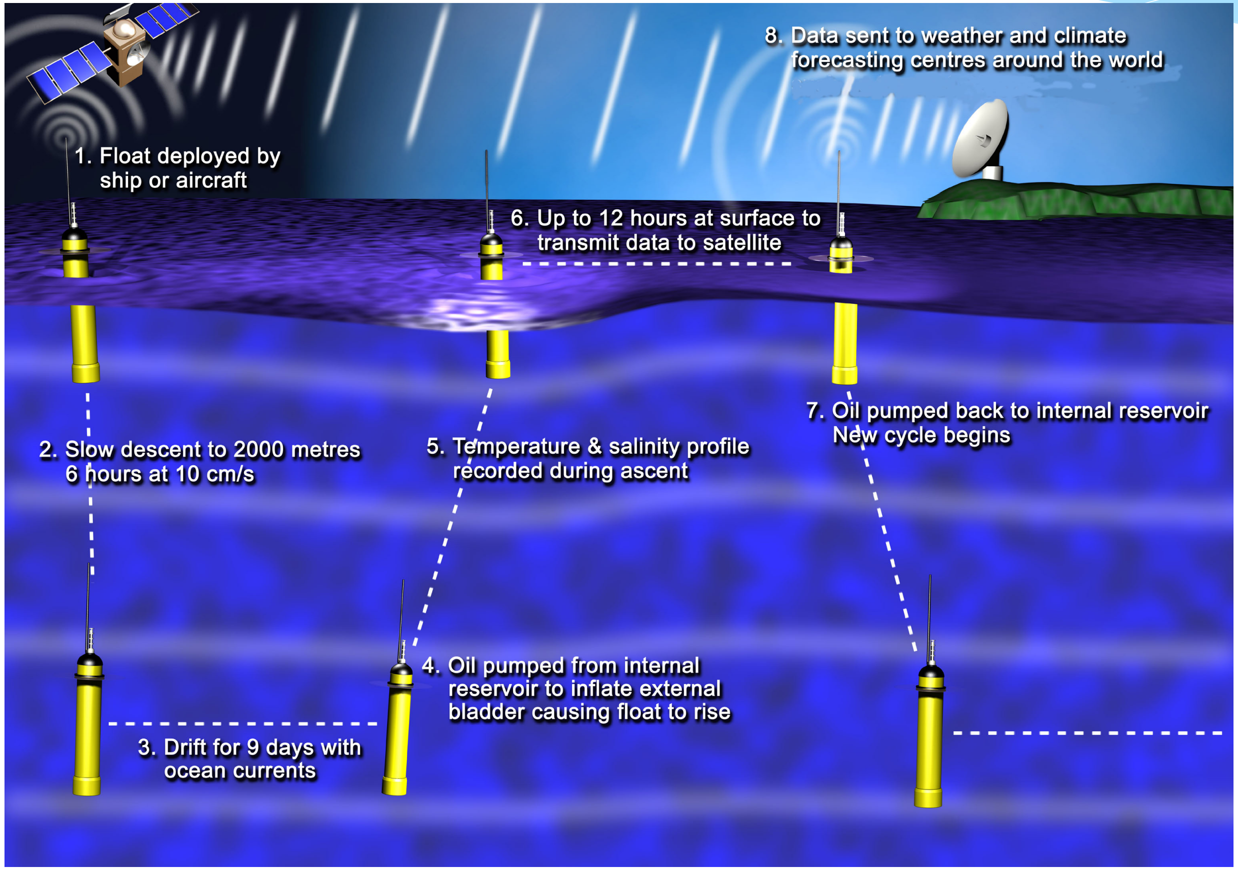
* No basic rule (e.g., Navier-Stokes equations) for biology
* Very nonlinear system
* Many uncertain and tunable parameters
* Biology sensitive to dynamics and dynamical instabilities
* Tracer concentrations are positive variables

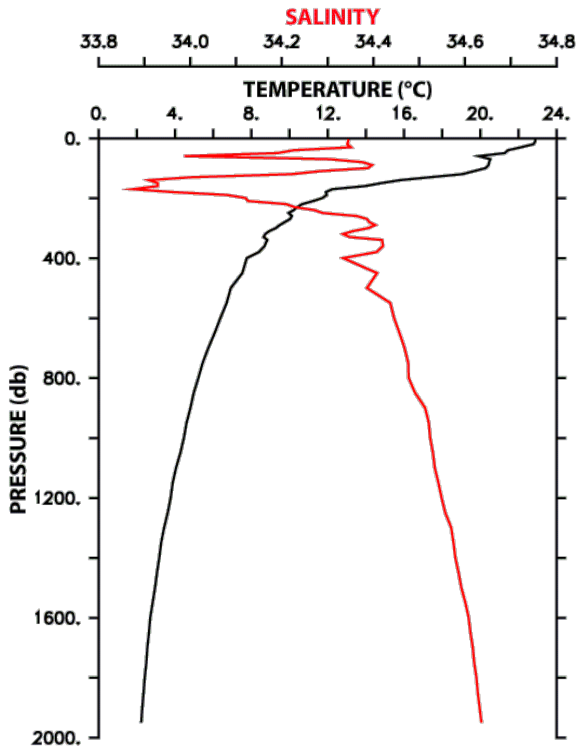
Observations

* In situ observations
  * Profiling floats: ARGO project
  * Moorings: OceanSITES project
  * Ships: SOOP and GOSUD projects, and WOCE program
  * Surface drifters: DBCP and E-SURFMAR projects
  * Gliders: EGO initiative
  * Marine mammals
* Satellite observations
  * Altimetry
  * Sea surface temperature (SST)
  * Ocean color

In situ observation #1: profiling floats

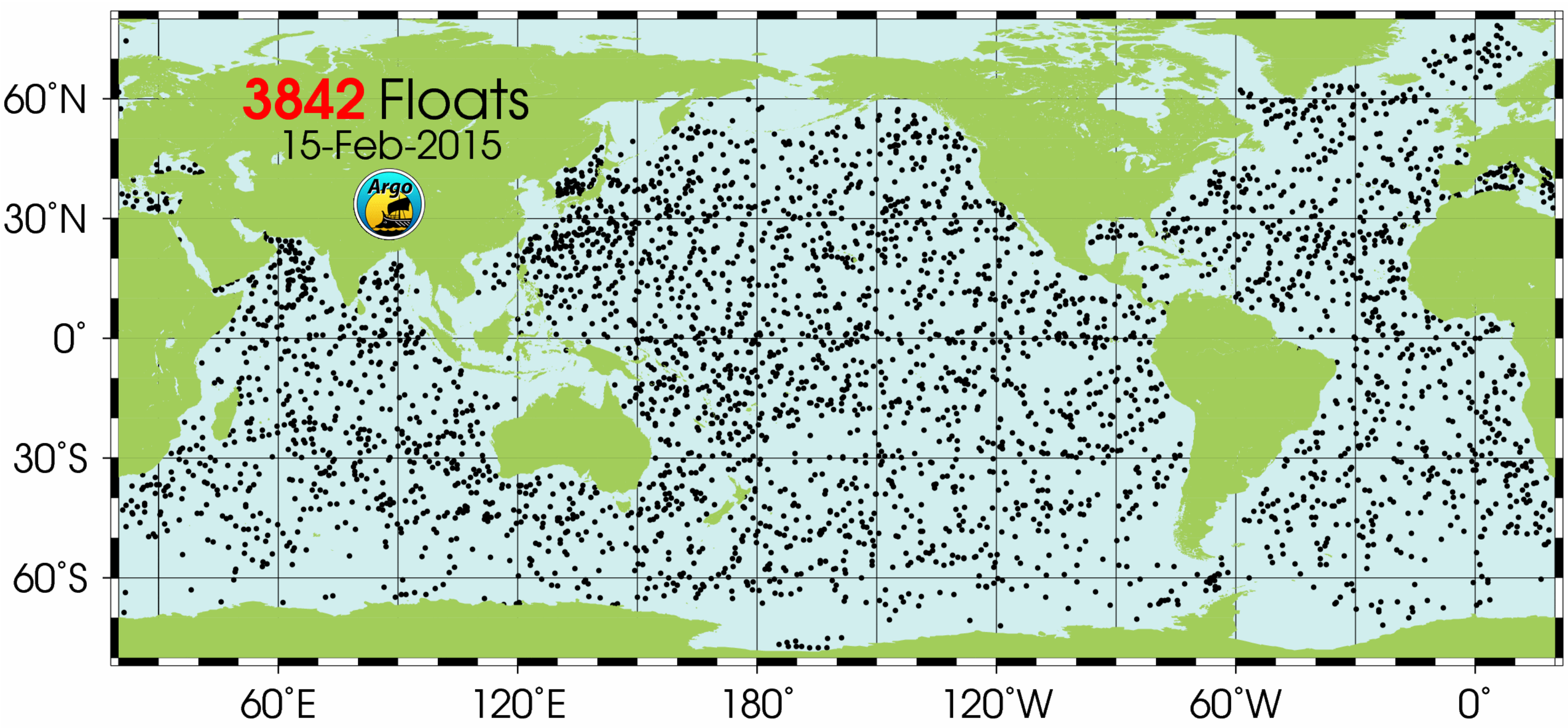
ARGO = network of profiling floats





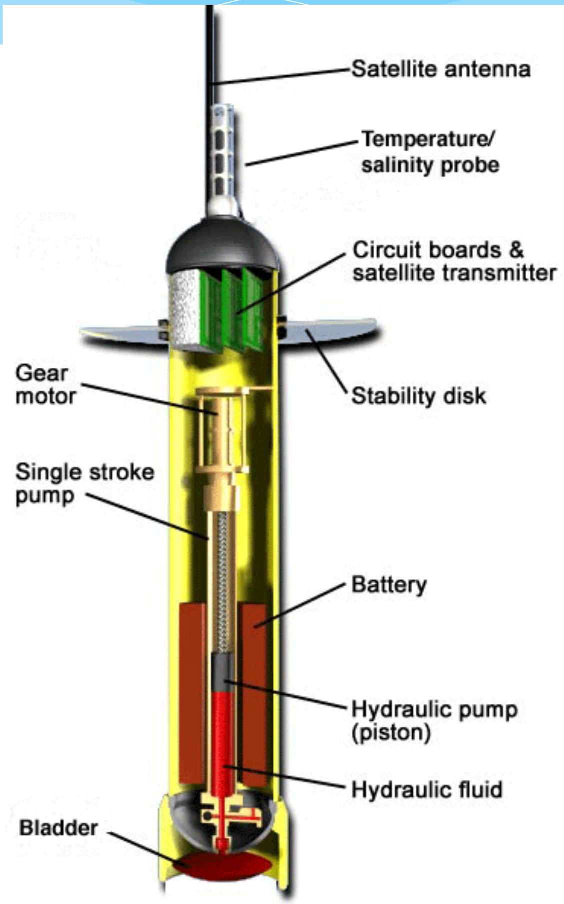
http://www.argo.ucsd.edu/

In situ observation #1: profiling floats



http://www.argo.ucsd.edu/

In situ observation #1: profiling floats



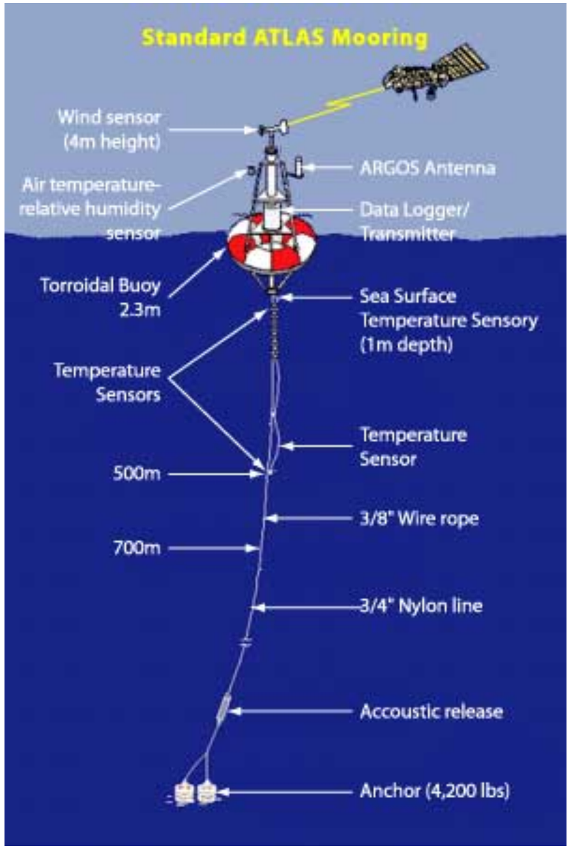
http://www.argo.ucsd.edu/

In situ observation #1: profiling floats

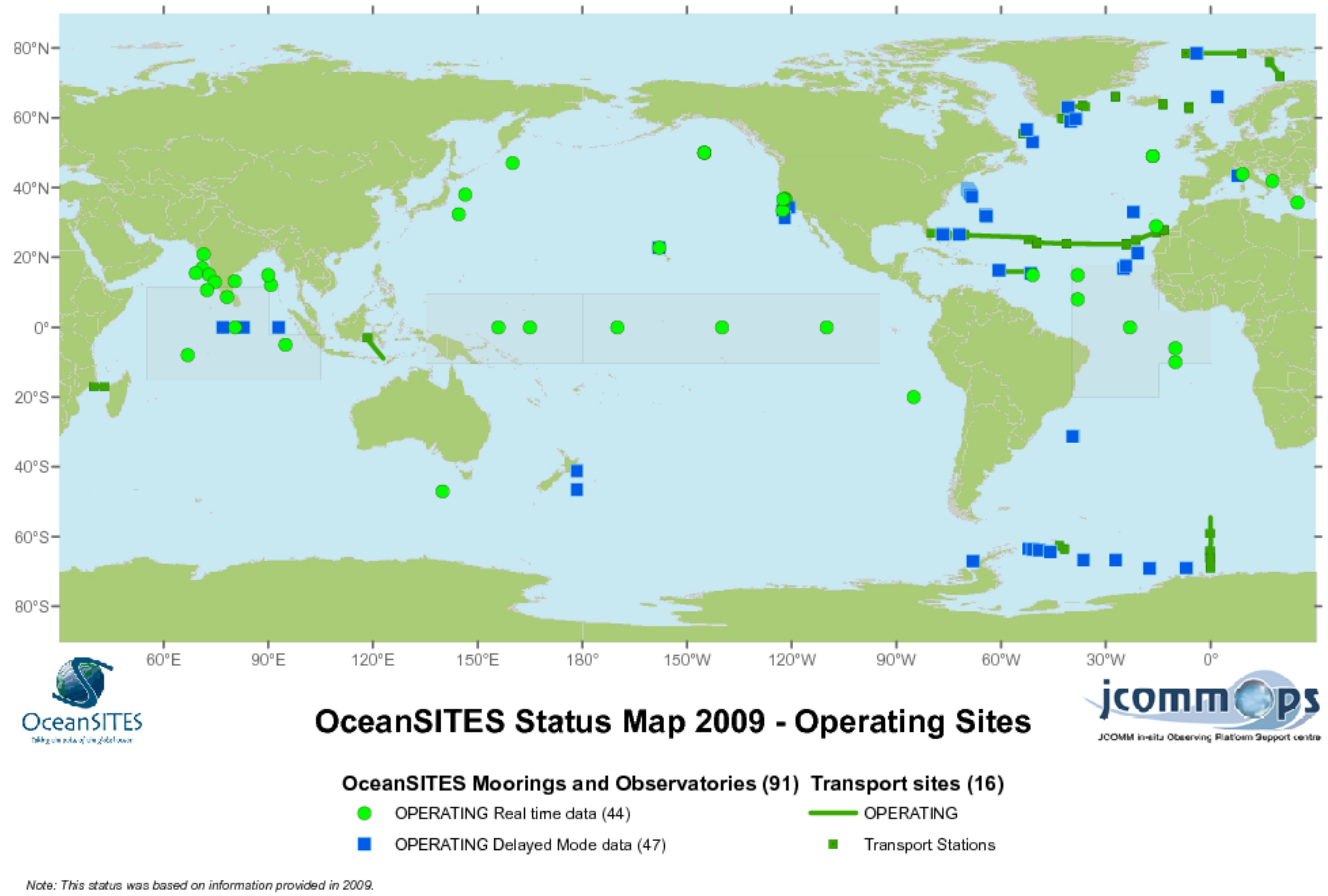
* +++ : Spatial coverage, vertical information, autonomy
* - - - : needs maintenance, some regions hard to sample, poor sampling

In situ observation #2: Moorings

Moorings are managed by the project OceanSITES.



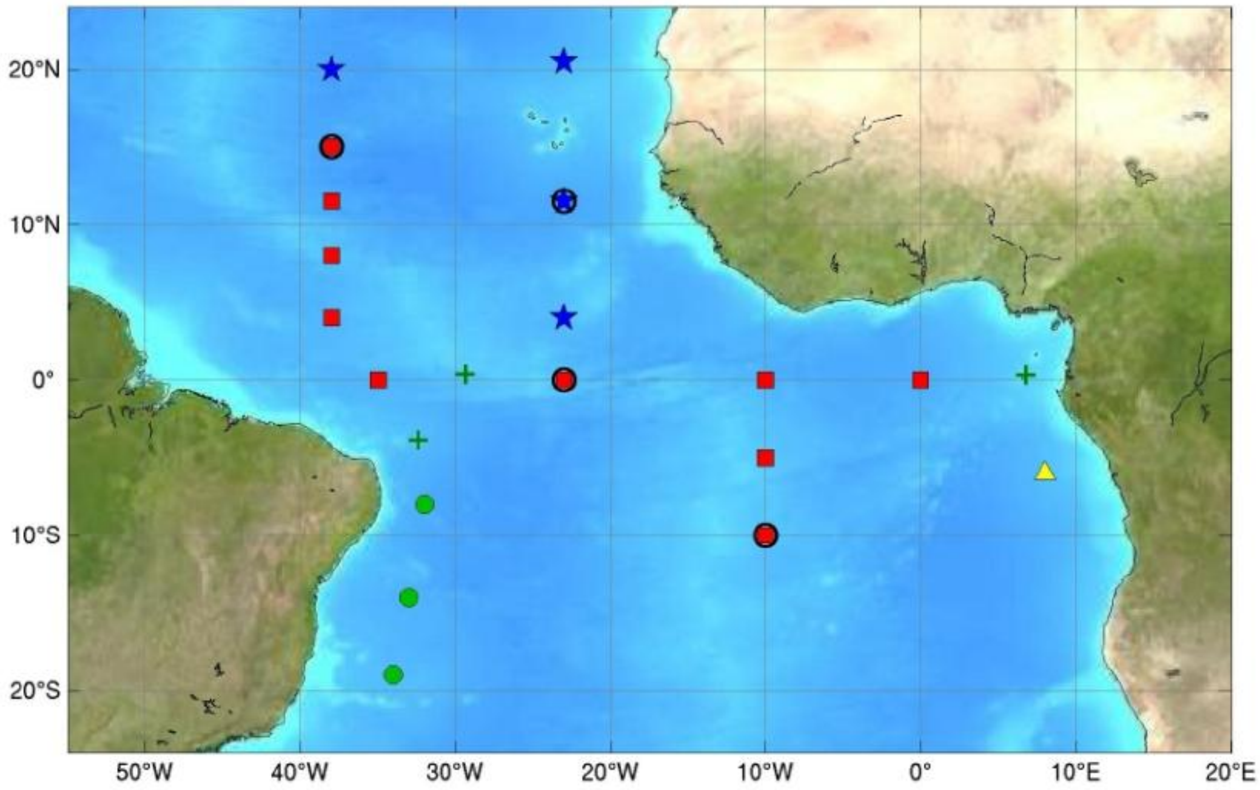
In situ observation #2: Moorings



http://www.whoi.edu/virtual/oceansites/network/index.html

In situ observation #2: Moorings

The French contribution to OceanSITES: PIRATA



http://www.brest.ird.fr/pirata/pirata.php

In situ observation #2: Moorings

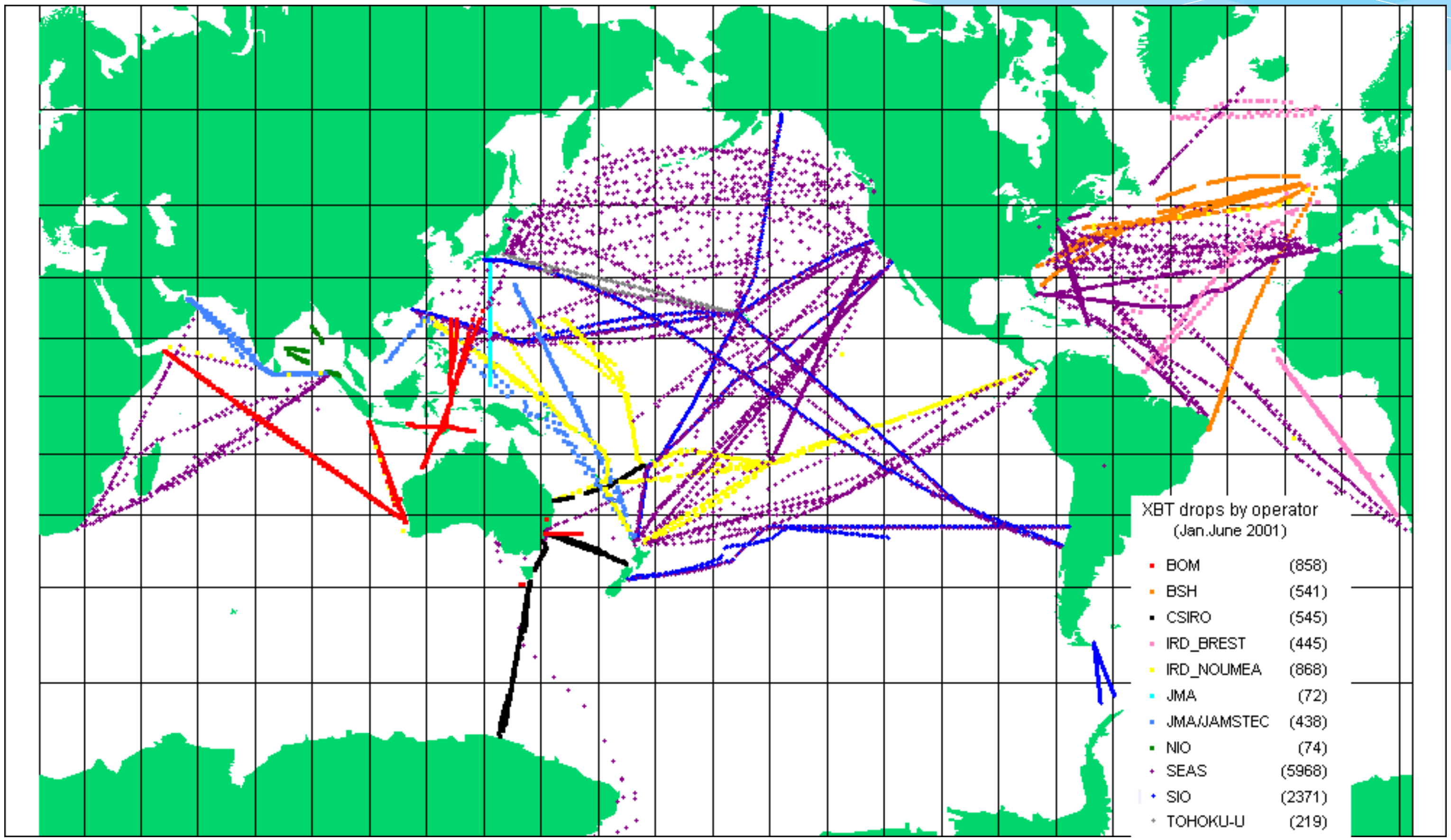
* +++ : time sampling, vertical information, autonomy
* - - - : expensive to build and maintain, poor spatial coverage

In situ observation #3: Ships

* Volunteer observing ships
* Research vessels

In situ observation #3: Ships

Ships of Opportunity: XBT drops

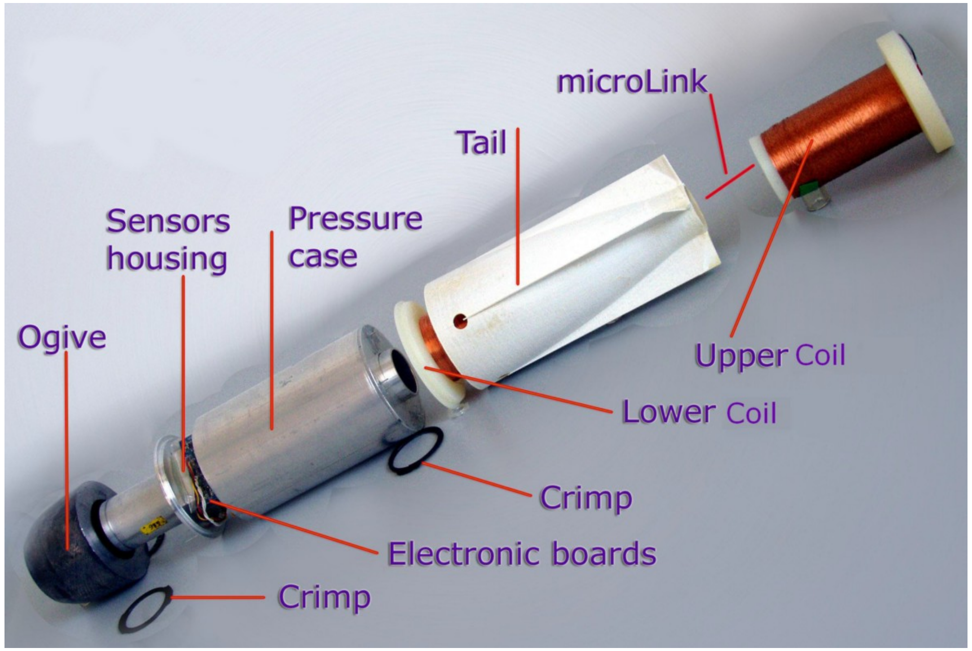


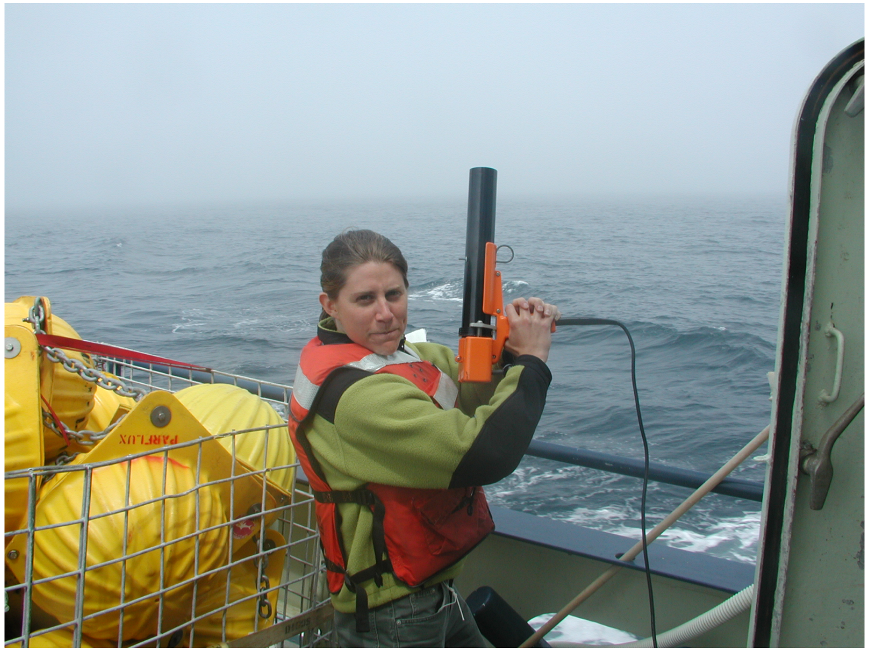
January-
June 2001

http://www.jcommops.org/soopip/

In situ observation #3: Ships

XBT probe, drop



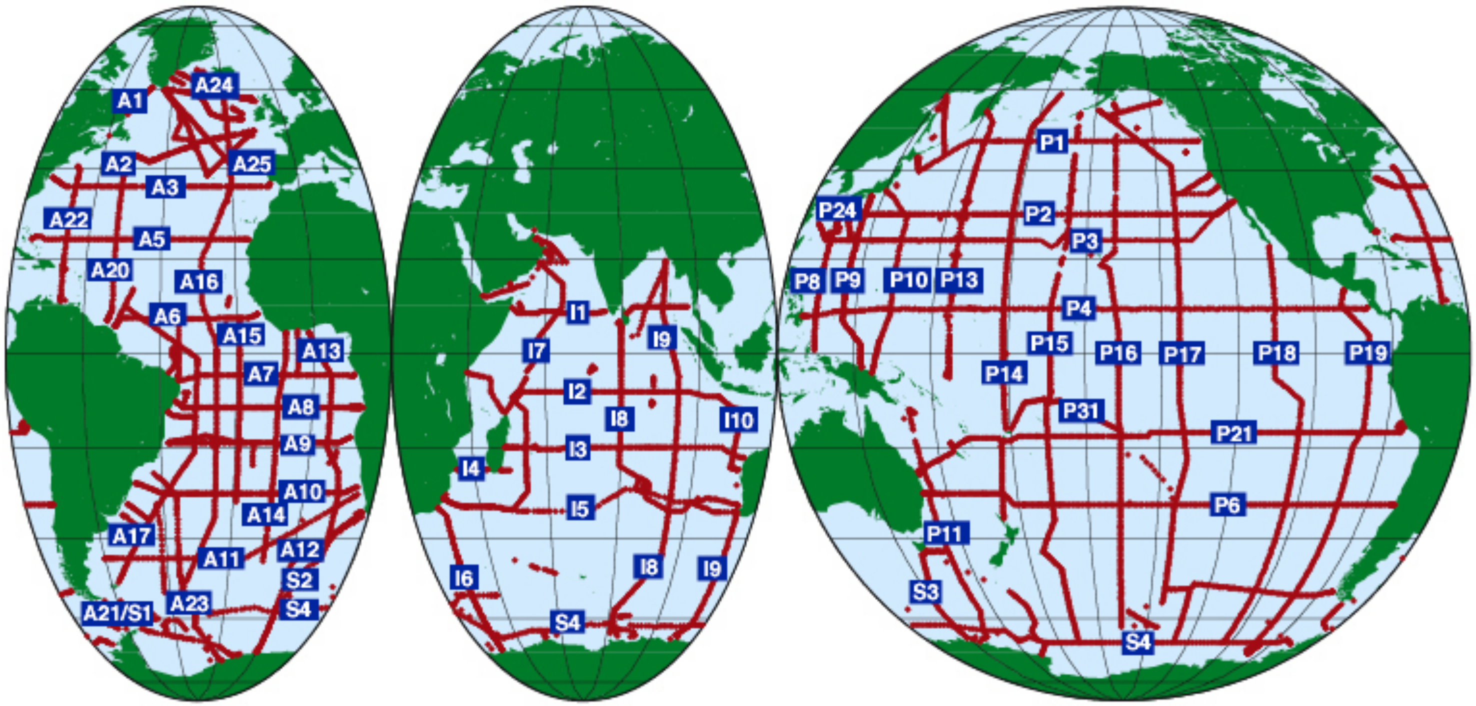


XBT: Expendable bathythermograph. Measure temperature and depth to ~1000 m

In situ observation #3: Ships

Research vessels: WOCE and isolated projects.

This survey took 10 years!



http://woceatlas.tamu.edu/

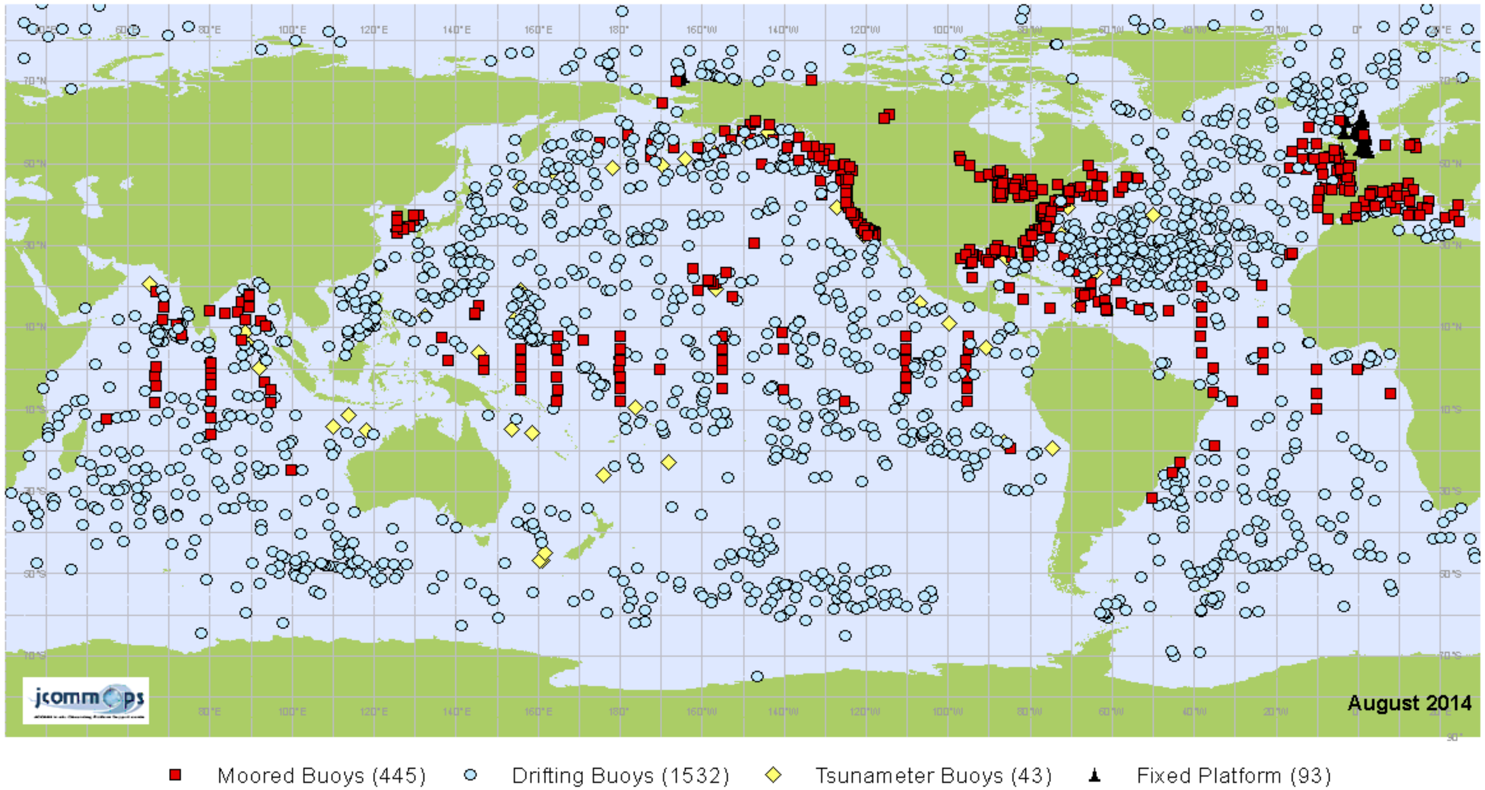
In situ observation #3: Ships

- VOS:
  - +++: cost effective, vertical information
  - - - - : limited to commercial routes, rarely deeper than 800 m
- Research vessels:
  - +++: often go to remote and poorly observed areas
  - - - - : extremely expensive, extremely poor coverage

Observations

In situ observation #4: surface drifters

Projects DBCP and E-SURFMAR



http://www.jcommops.org/dbcp/

In situ observation #4: surface drifters

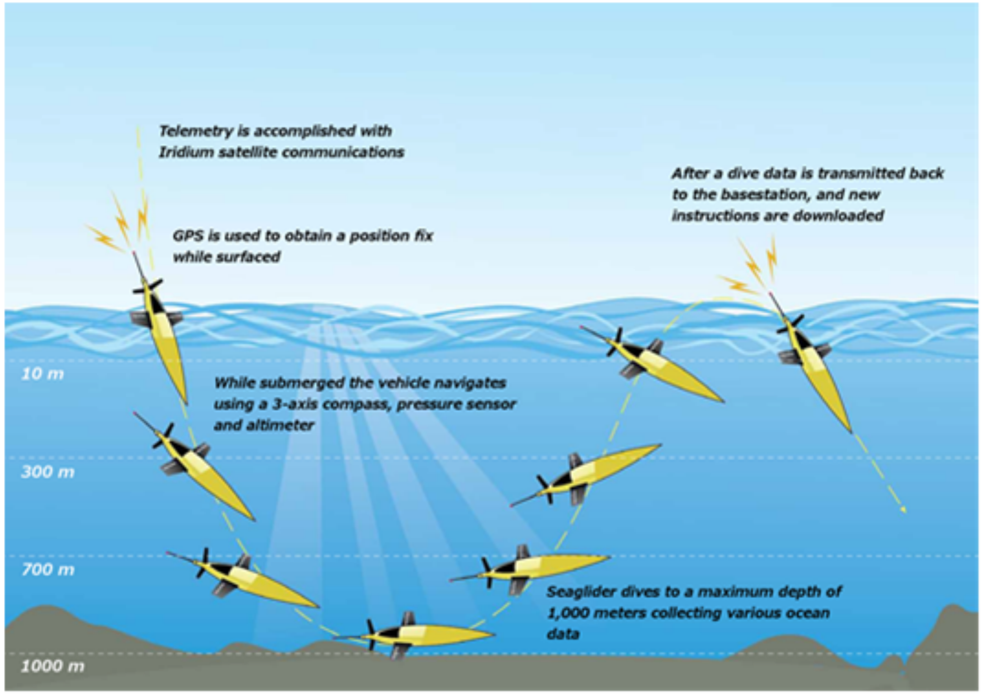
A drifter measures surface temperature and currents.
http://www.aoml.noaa.gov/
http://www.nefsc.noaa.gov/

In situ observation #4: surface drifters

* +++ : Spatial coverage, autonomy
* - - - : needs maintenance, some regions hard to sample, poor sampling

In situ observation #5: gliders

Gliders organized within the EGO initiative

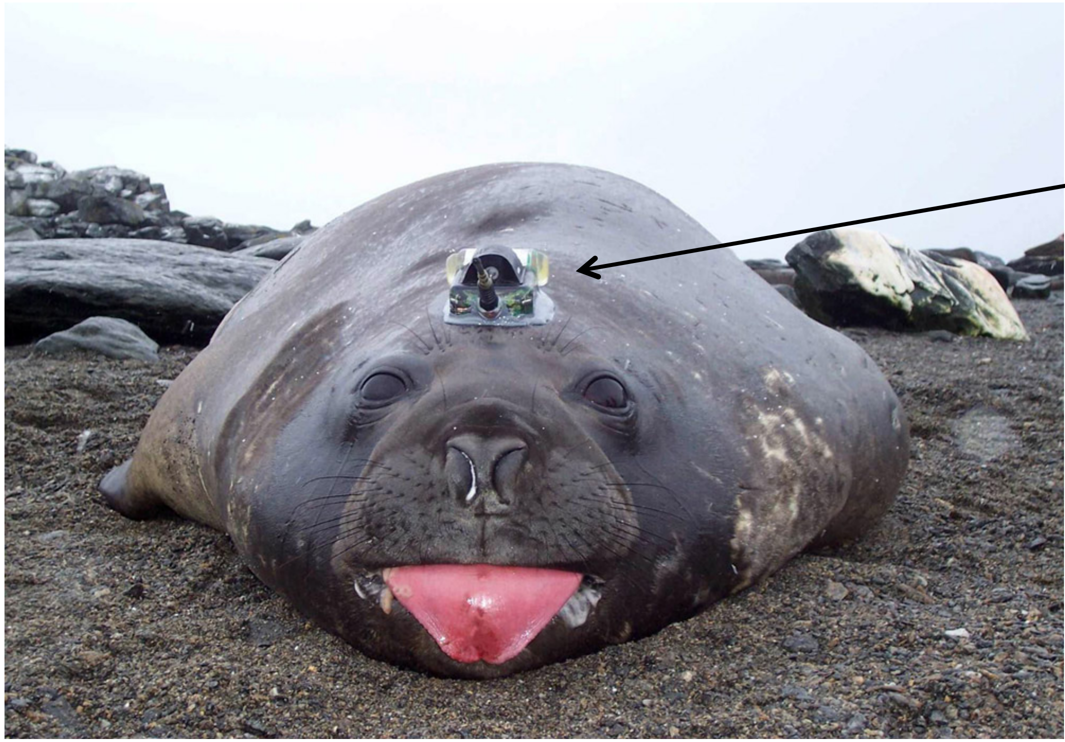


http://www.fastwave.com.au

In situ observation #5: gliders

* +++: flexible, vertical information
* - - - : limited to targeted campaigns

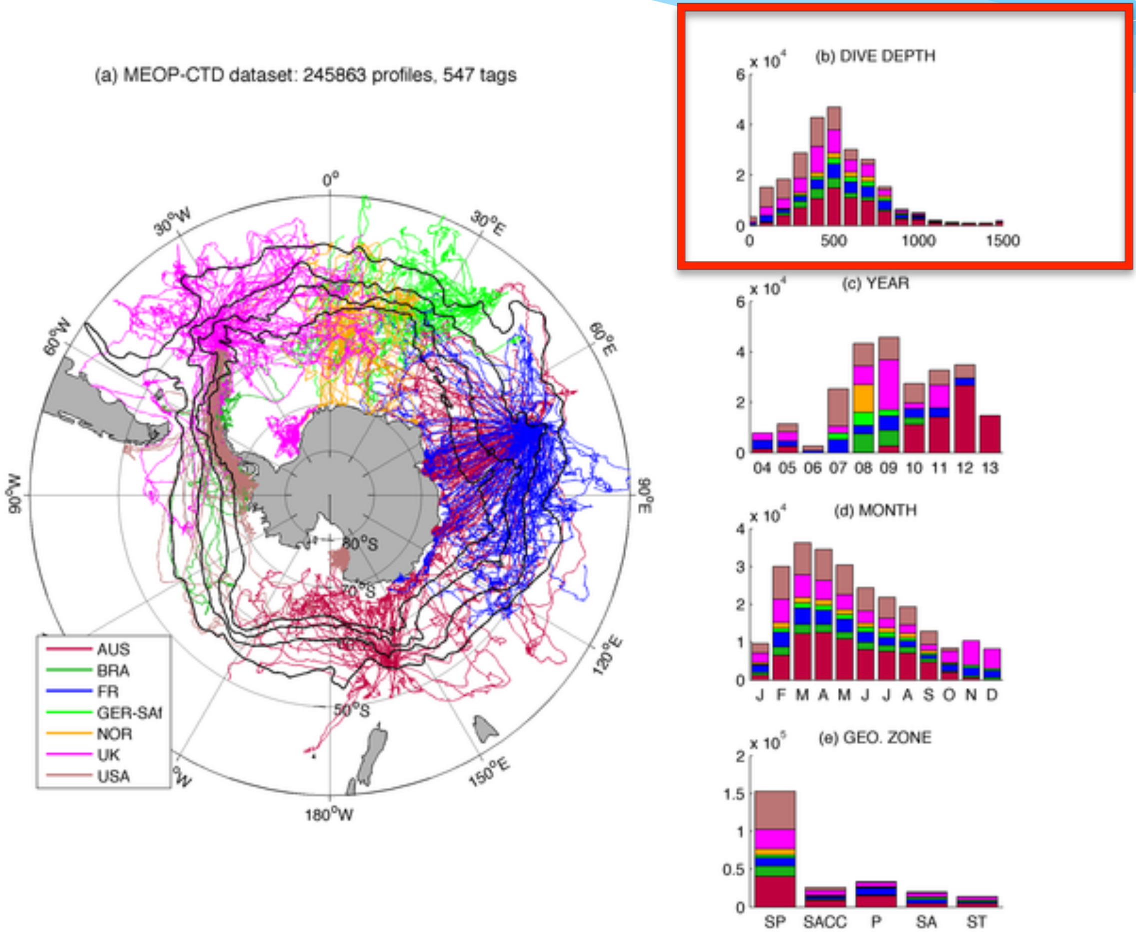
In situ observation #6: marine mammals



A miniaturized CTD (Conductivity-Temperature-Depth) probe

Observations

In situ observation #6: marine mammals

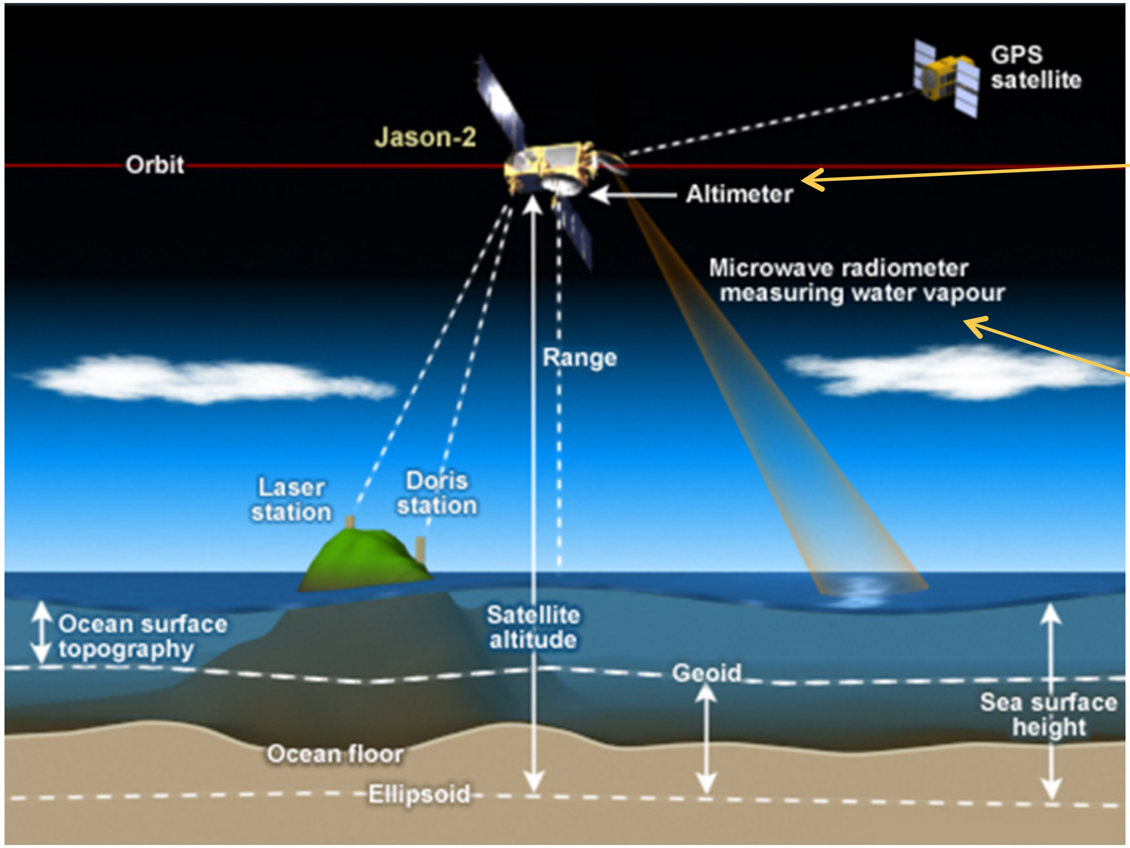


Sample poorly observed areas!

In situ observation #6: marine mammals

* +++: access to poorly observed area, vertical information
* - - - : limited spatial and temporal coverage

Satellite observation #1: altimetry



Radar altimeter (emitter & antenna)

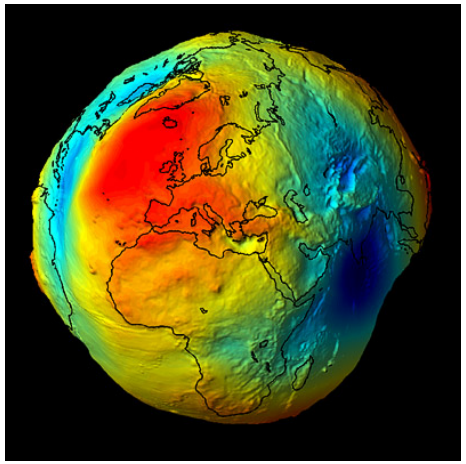
For atmospheric corrections

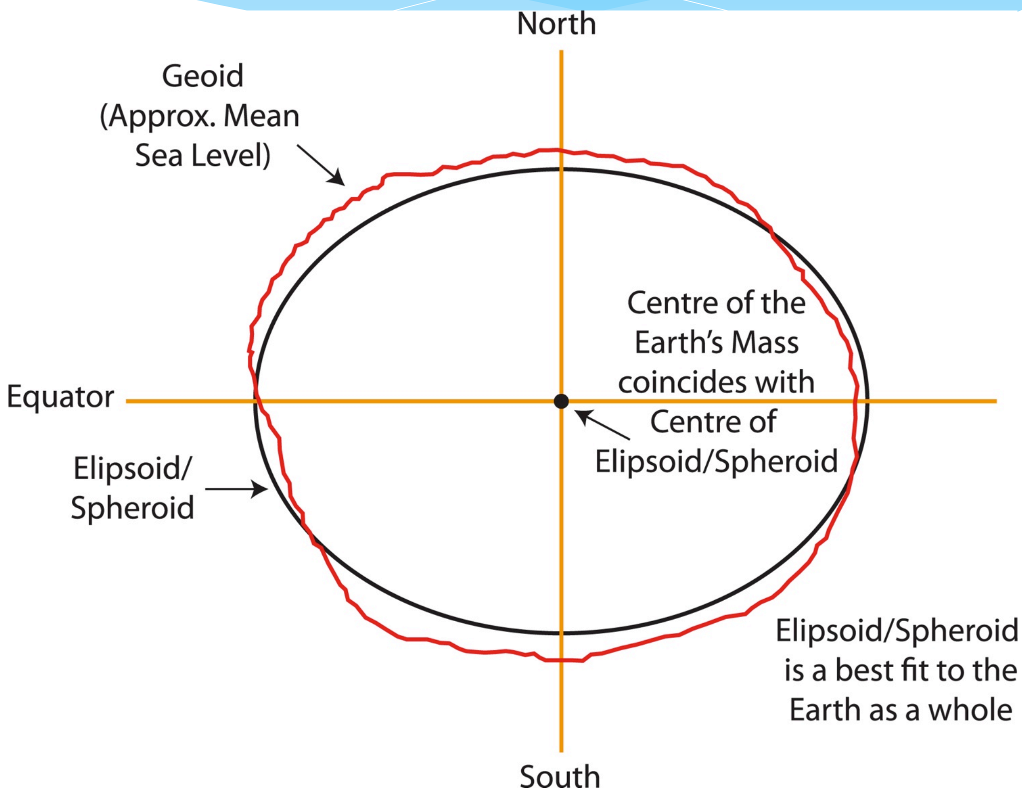
Height of the satellite: ~1340 km

Satellite observation #1: altimetry

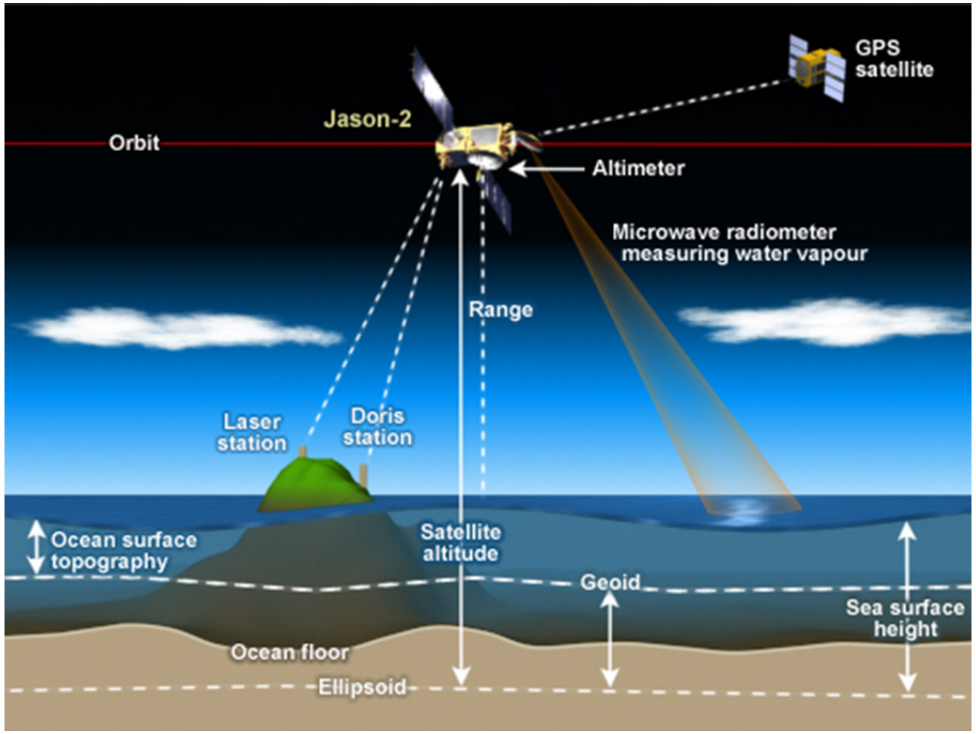
Ellipsoid: theoretical ellipsoidal surface matching approximately the shape of the Earth at sea level.

Geoid: equipotential surface of the effective gravitational field of the Earth at mean sea level.





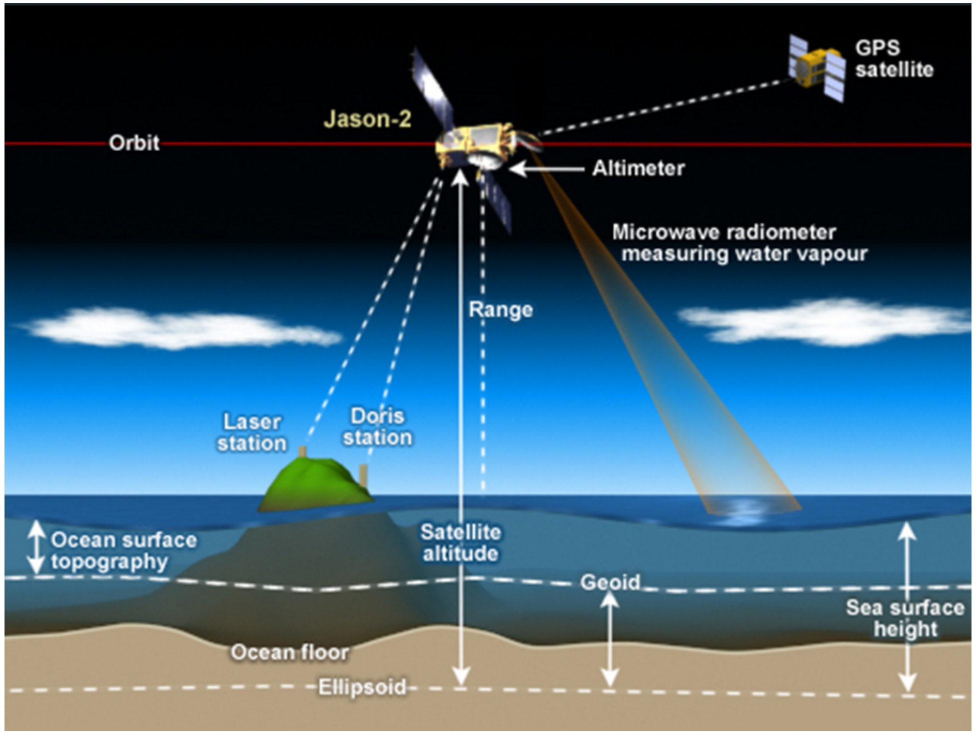
Satellite observation #1: altimetry



The altimeter measures the distance to the surface, R_{alt}, and its own altitude, H_{alt} (ref. ellipsoid). From them we get the sea surface height (SSH):

$$SSH = H_{alt} - R_{alt}$$

Satellite observation #1: altimetry

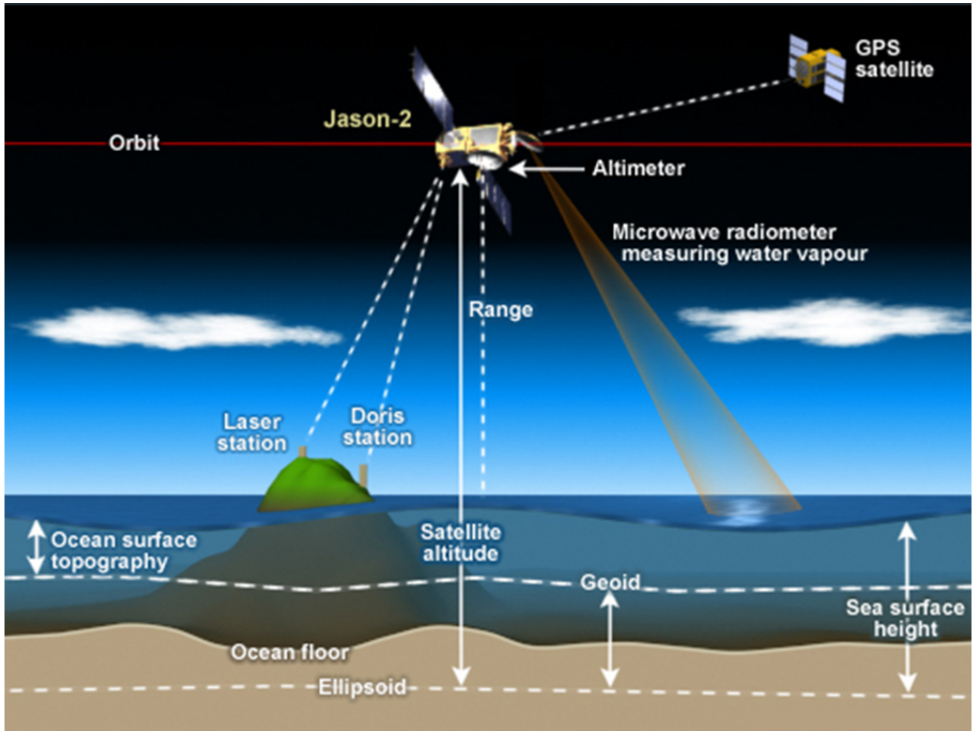


$SSH = H_{alt} - R_{alt}$

The sea surface height (SSH) gathers contributions from gravity (geoid) and ocean surface topography (tides, atmospheric pressure, and ocean dynamics):

$SSH = h_{geoid} + h_{tide} + h_{atm} + h_{dyn}$

Satellite observation #1: altimetry



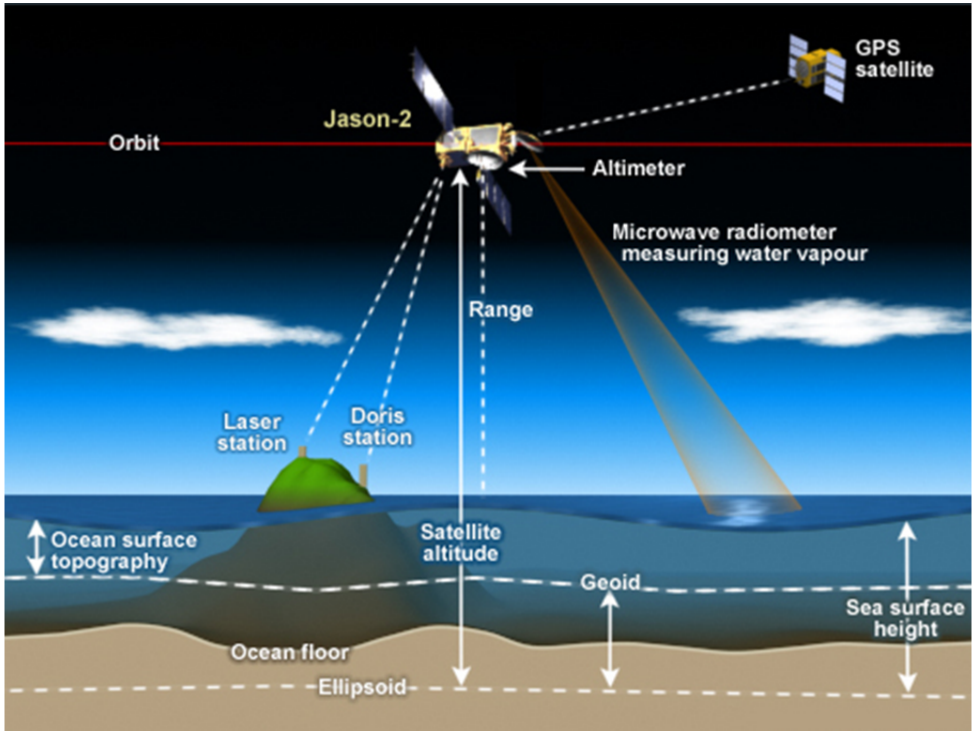
$SSH = h_{geoid} + h_{tide} + h_{atm} + h_{dyn}$

h_{dyn} is the ocean dynamic topography and is due to the motions of the sea. It is the relevant term for studying the ocean circulation.

$h_{dyn=} = H_{alt} - R_{alt} - h_{geoid} - h_{tide} - h_{atm}$

The accuracy of h_{dyn} estimation does not depend only on the altimetric measurement itself.

Satellite observation #1: altimetry



$h_{dyn=}=H_{alt}-R_{alt}-h_{geoid}-h_{tide}-h_{atm}$

Precisions:
H_{alt}: 2 cm
h_{tide}: 2 cm

h_{atm} correction is based on atmospheric models.

h_{geoid} is the big issue.

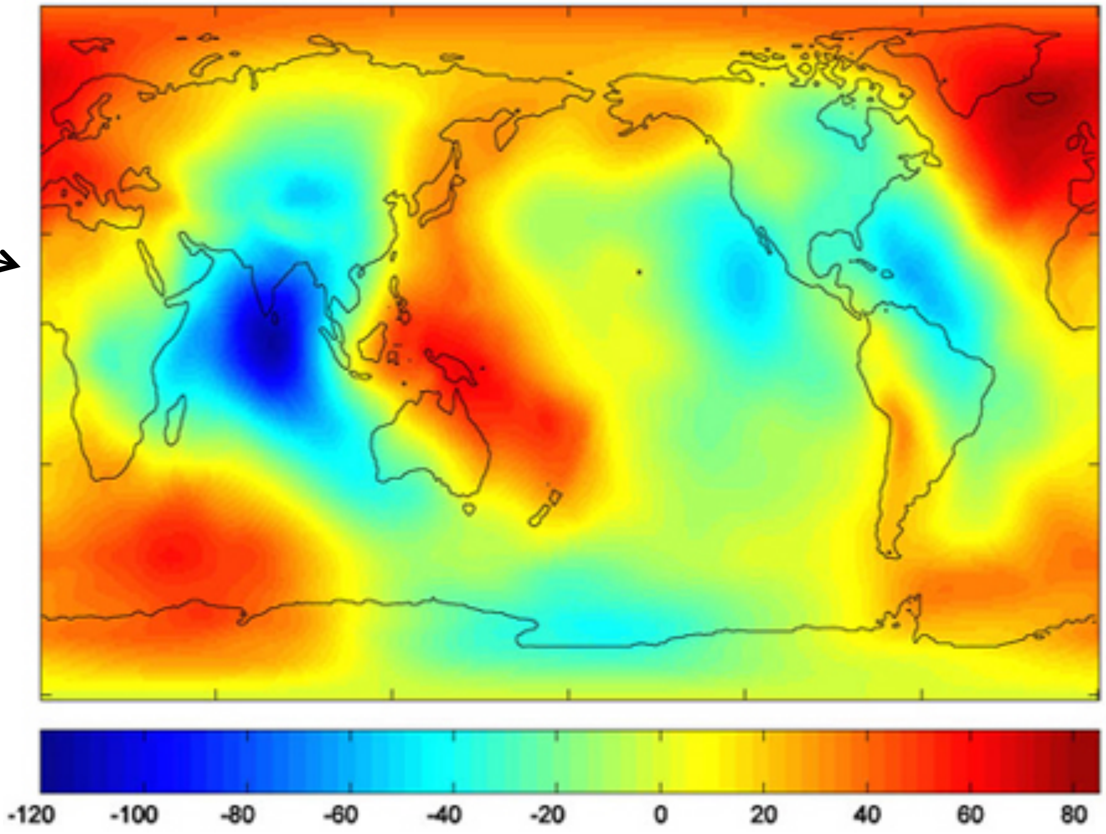
Observations

Satellite observation #1: altimetry

Recent satellite missions provides a geoid: GOCE, GRACE.

But so far we keep h_{dyn} $+h_{geoid}$ and substract the time mean over many orbit cycles to obtain the sea level anomaly (SLA).

SLA does not contain any information about the mean ocean circulation.



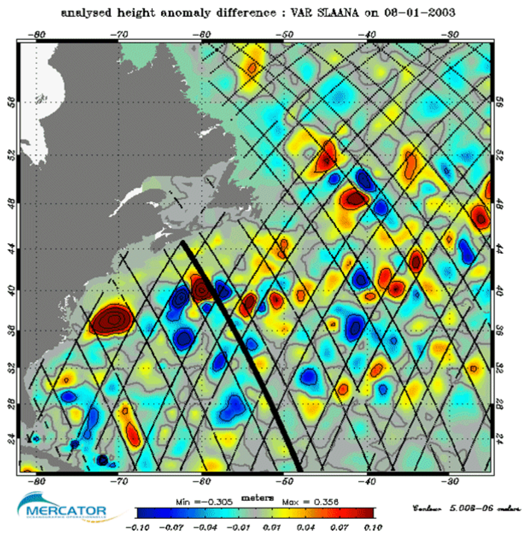
(University of Texas Center for Space Research and NASA)

Satellite observation #1: altimetry

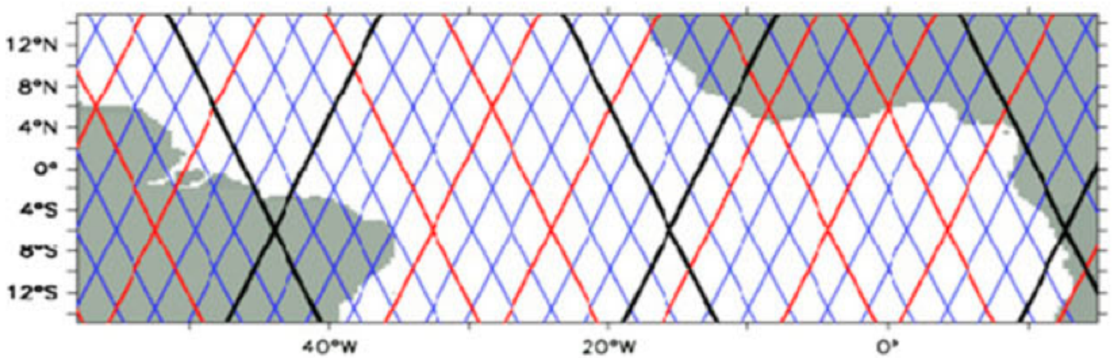
Obtaining a "good" SLA products usable in ocean DA systems requires a complex data post-processing, from quality control to calibration…

Since recently this was performed by the company CLS (Collecte Localisation Satellites), affiliate of CNES and IFREMER, within the project AVISO. It is now operated by the CMEMS.

Satellite observation #1: altimetry



Orbit of Jason:
Cycle of 10 days.



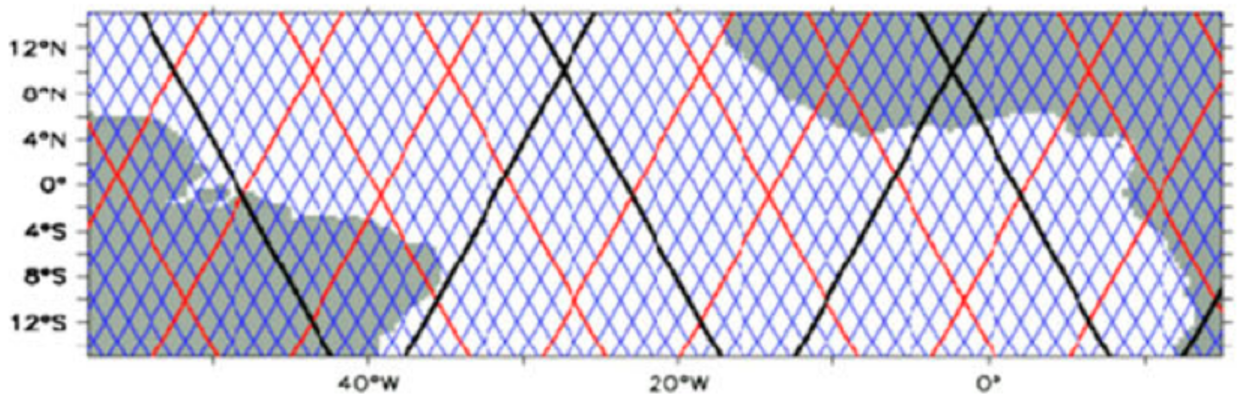
Orbit-1 (Jason)

H=1336km i=66°

(sub-)cycles (days) : **0.9** 3.3 9.9

Satellite observation #1: altimetry

Orbit of GFO:
Cycle of 17 days.



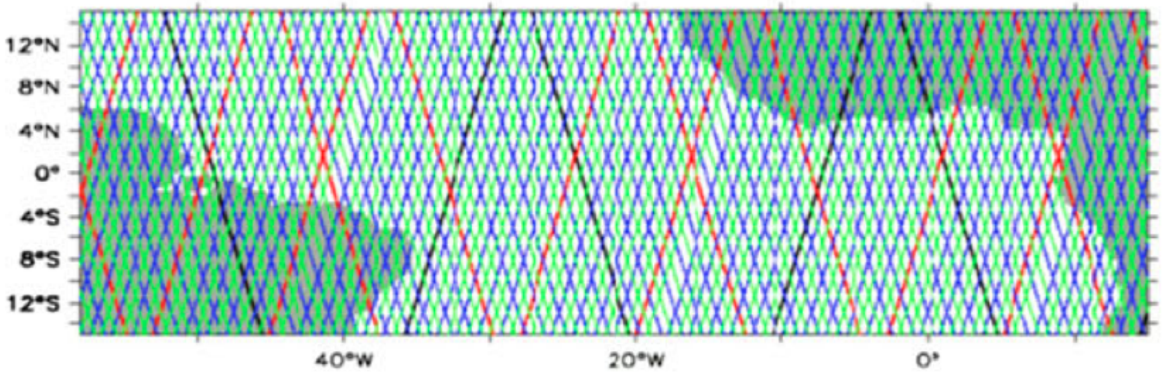
Orbit-2 (Gfo) : H=800km i=108°

H=800km i=108°

(sub-)cycles (days) : **1.0** 2.8 17.0

Satellite observation #1: altimetry

Orbit of Envisat and Saral:
Cycle of 35 days

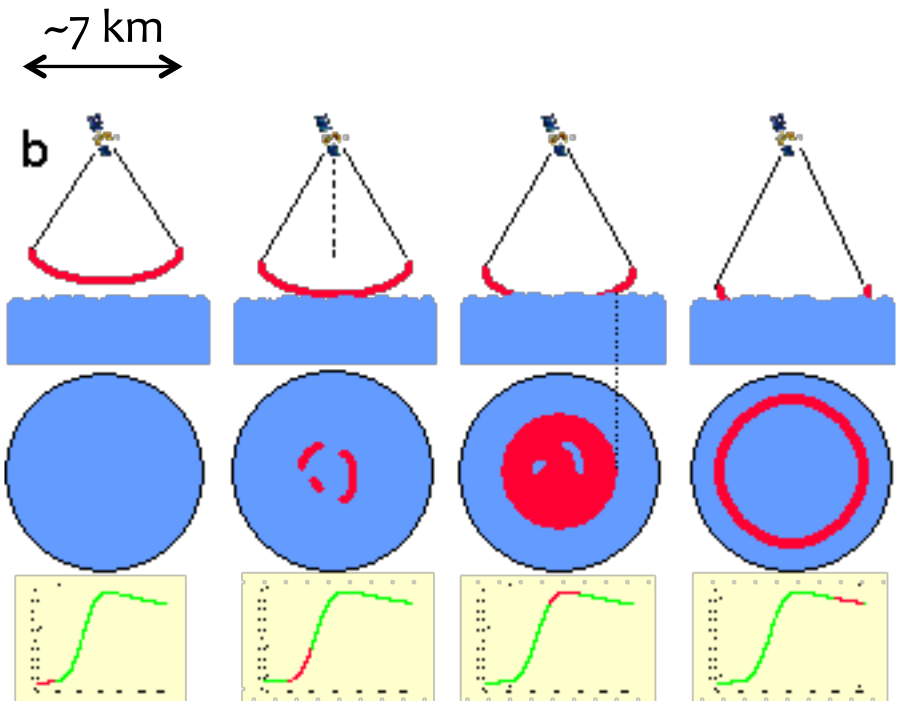


Orbit-3 (Envisat, Saral

H=782km i=98°

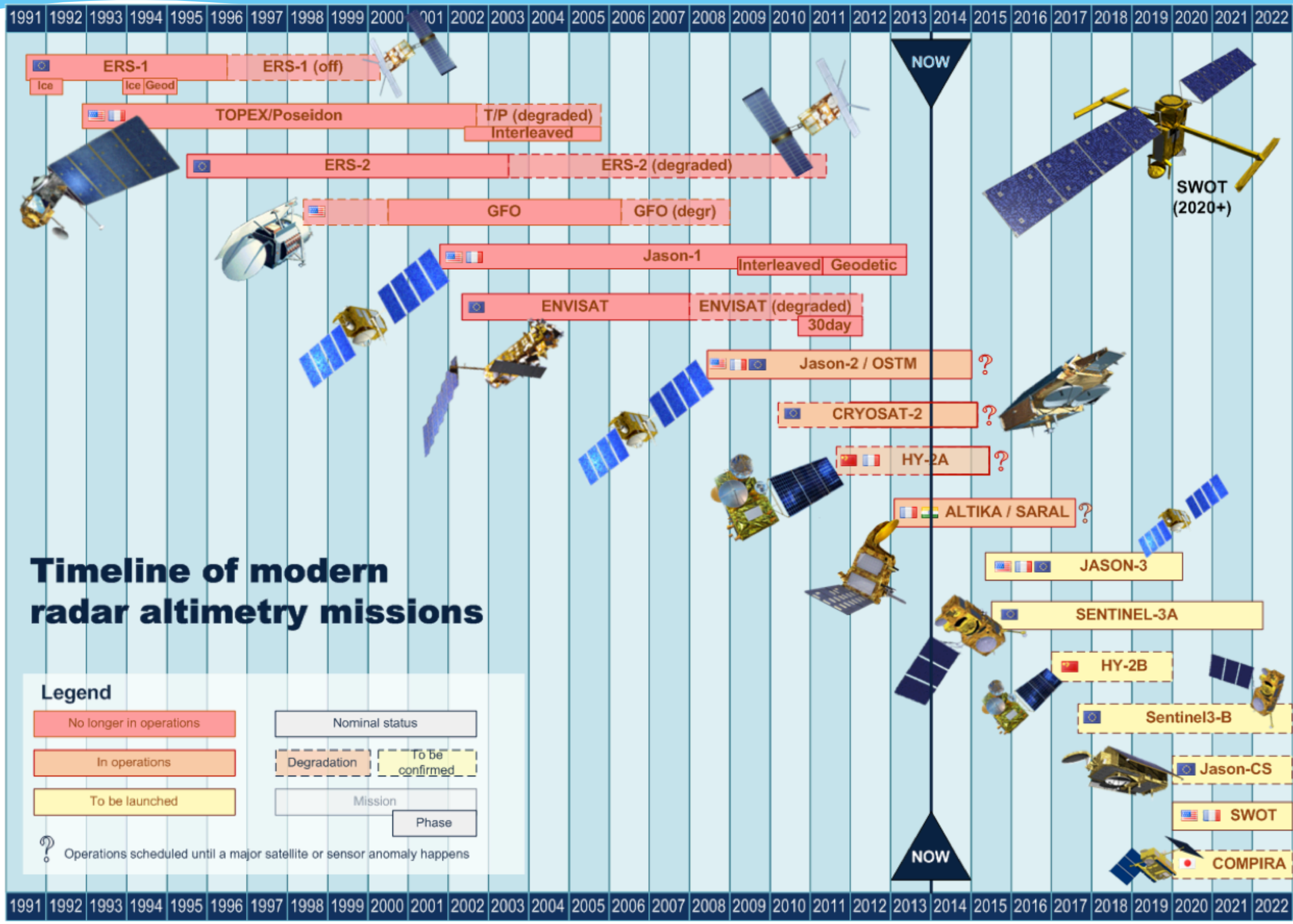
(sub-)cycles (days) : **1.0** 3.0 17.5 35.0

Satellite observation #1: altimetry

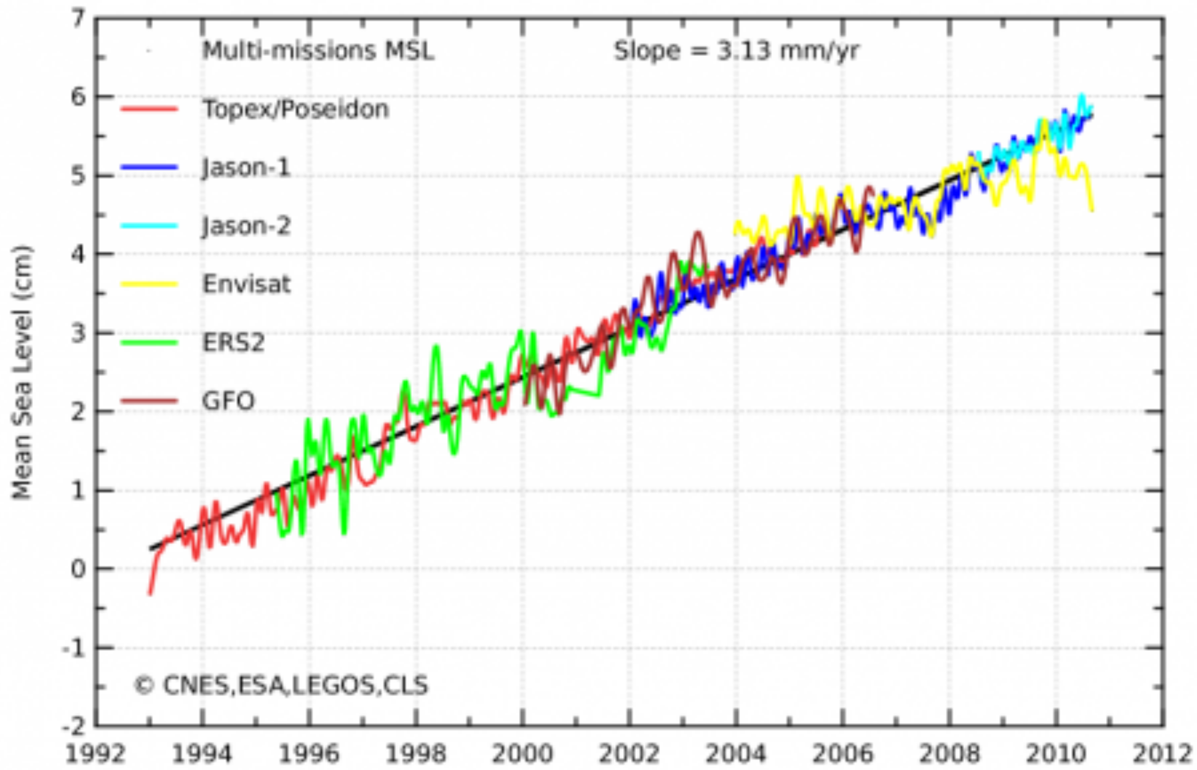


Radar altimetry provides information about mesoscale ocean topography (50-100 km) and waves.

Satellite observation #1: altimetry



Satellite observation #1: altimetry



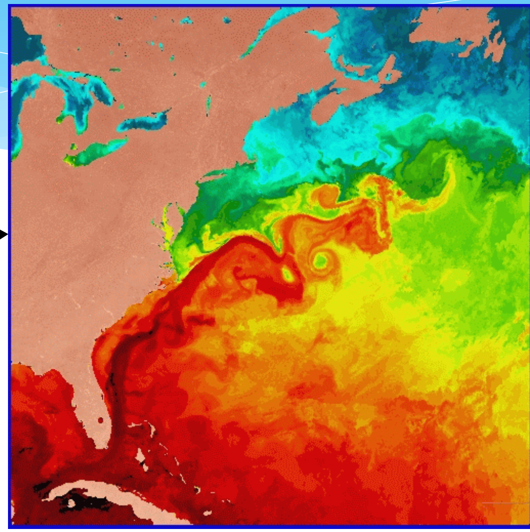
The continuity of satellite altimeters is essential for monitoring the mean sea level.

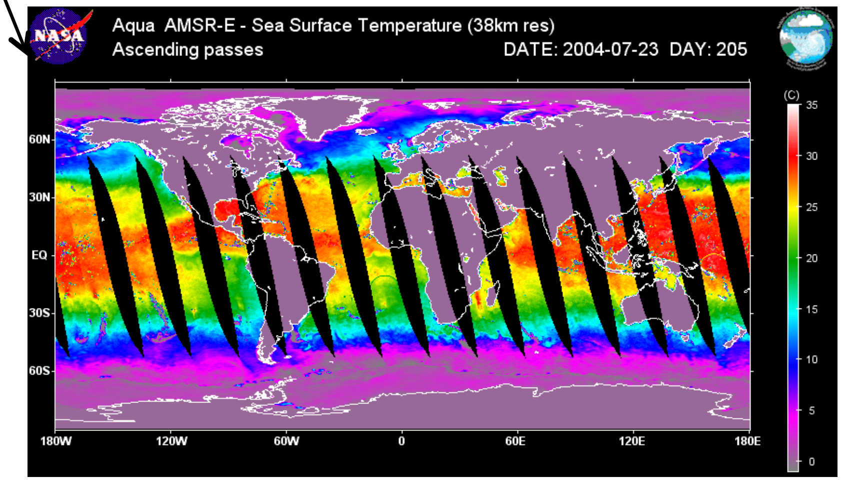
Satellite observation #2: SST

- IR radiometer (e.g. AVHRR)
- Microwave radiometer (e.g. AMSR-E)

- Both at 1-km resolution.
- MW insensitive to clouds but less sensitive and easy to calibrate.

Some IR sensors are on-board geostationary satellites (res. 5 km). Most are polar orbiting.





Satellite observation #2: SST

IR sensors records the radiance detected at the top of the atmosphere in various bands, linked to temperature by the Planck equation (black body emission):

$$L(\lambda, T) = \frac{C_1}{\pi\lambda^5[\exp(C_2/\lambda T) - 1]}$$

The spectral radiance is transformed into a brightness temperature after direct calibration of the sensors using on-board black-body targets.

Atmospheric corrections are derived from the combination of the signals in different spectral bands. Calibration is based on in-situ measurements of SST. Alternative approach: explicit simulation of the atmospheric radiative transfer.

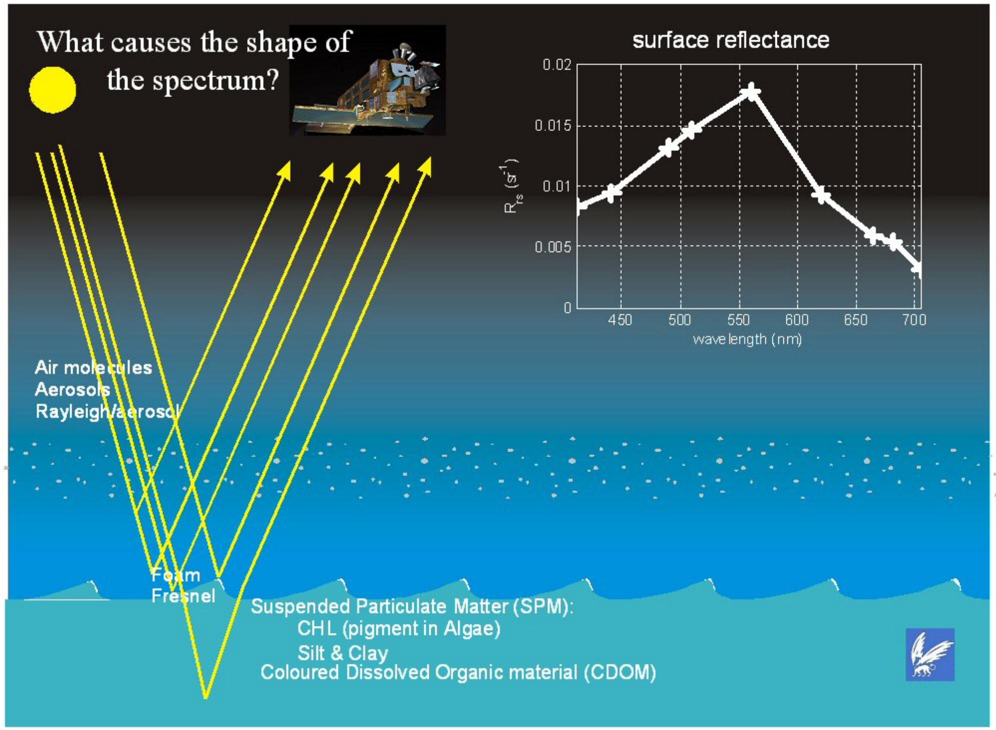
Satellite observation #2: SST

Two issues with satellite SST from the DA viewpoint:

- Cloud detection
- SST is a “skin” temperature (representation error)

Satellite observation #3: Ocean color

Ocean color sensors record reflectances in the solar spectrum.

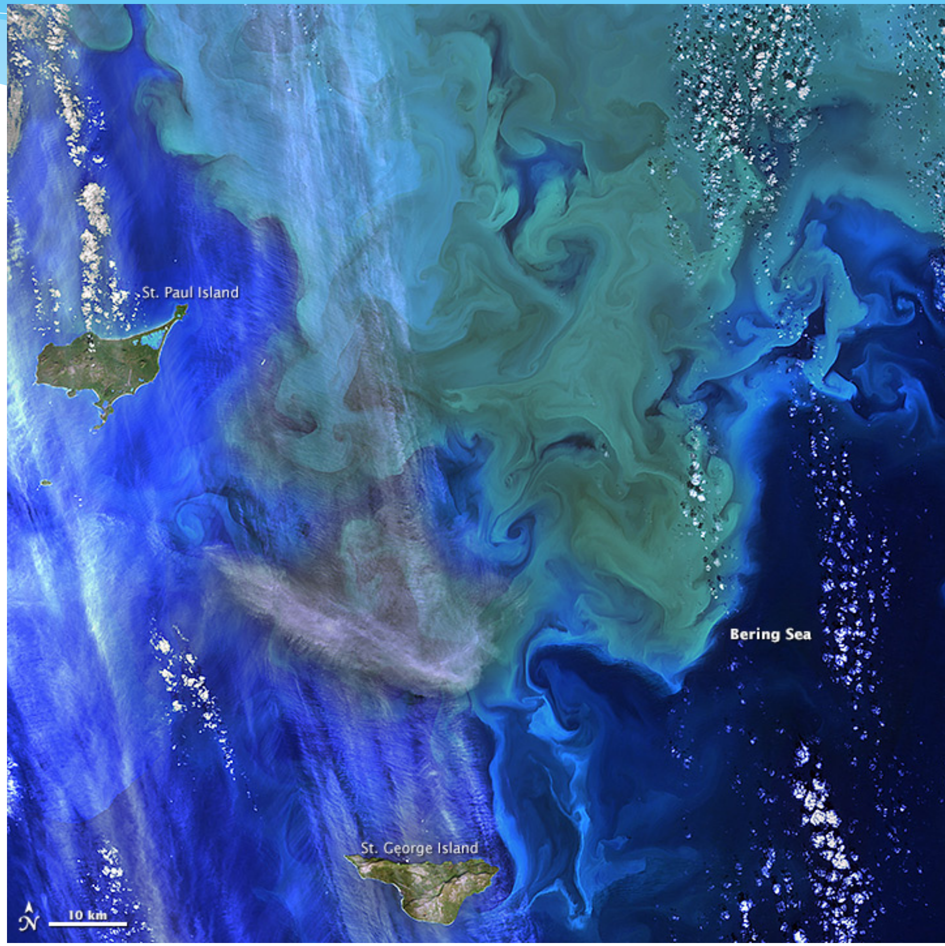


http://www.seos-project.eu/

Satellite observation #3: Ocean color

Ocean color sensors detect chlorophyll.

Left: A phytoplankton bloom captured near Alaska by Operational Land Imager (OLI) on Landsat 8 (NASA).



Satellite observation #3: Ocean color

Proof of concept: CZCS (Coastal Zone Color Scanner), 1978-1986.
First operational ocean color products: SeaWIFS (Sea-viewing Wide Field-of-view Sensor), 1997-2010

In addition to the various measurement errors (atmospheric corrections, etc), a significant source of error lies in the algorithm to retrieve chlorophyll concentrations. The accepted error is 30% in general.

Observations: summary

* Quite large diversity of in situ data, but rather sparse;
* A large amount of satellite data, but satellites only see the surface;
* They all contain uncertainties (measurement or representation) that are difficult to estimate.

Ocean DA using Ensemble Kalman filters

* Ensemble Kalman filters
* Localization
* Incremental Analysis Updating (IAU)
* Bogus
* Gaussian anamorphosis
* About the observation error covariance matrix

Ensemble Kalman filters

Kalman filter equations:

Initialization: $\mathbf{x}_0^f$ and $\mathbf{P}_0^f$

Analysis step:

$$
\begin{aligned}
\mathbf{K}_k &= (\mathbf{H}_k\mathbf{P}_k^f)^T[\mathbf{H}_k(\mathbf{H}_k\mathbf{P}_k^f)^T + \mathbf{R}_k]^{-1}, \\
\mathbf{x}_k^a &= \mathbf{x}_k^f + \mathbf{K}_k(\mathbf{y}_k^o - \mathbf{H}_k\mathbf{x}_k^f), \\
\mathbf{P}_k^a &= (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{P}_k^f.
\end{aligned}
$$

Forecast step:

$$
\begin{aligned}
\mathbf{x}_{k+1}^f &= \mathbf{M}_{k,k+1}\mathbf{x}_k^a, \\
\mathbf{P}_{k+1}^f &= \mathbf{M}_{k,k+1}\mathbf{P}_k^a\mathbf{M}_{k,k+1}^T + \mathbf{Q}_k.
\end{aligned}
$$

Ensemble Kalman filters

Kalman filter equations:

Initialization: $\mathbf{x}_0^f$ and $\mathbf{P}_0^f$

Analysis step:

Too big to store

$$\begin{aligned} \mathbf{K}_k &= (\mathbf{H}_k\mathbf{P}_k^f)^T[\mathbf{H}_k(\mathbf{H}_k\mathbf{P}_k^f)^T + \mathbf{R}_k]^{-1}, \\ \mathbf{x}_k^a &= \mathbf{x}_k^f + \mathbf{K}_k(\mathbf{y}_k^o - \mathbf{H}_k\mathbf{x}_k^f), \\ \mathbf{P}_k^a &= (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{P}_k^f. \end{aligned}$$

Often too big to invert

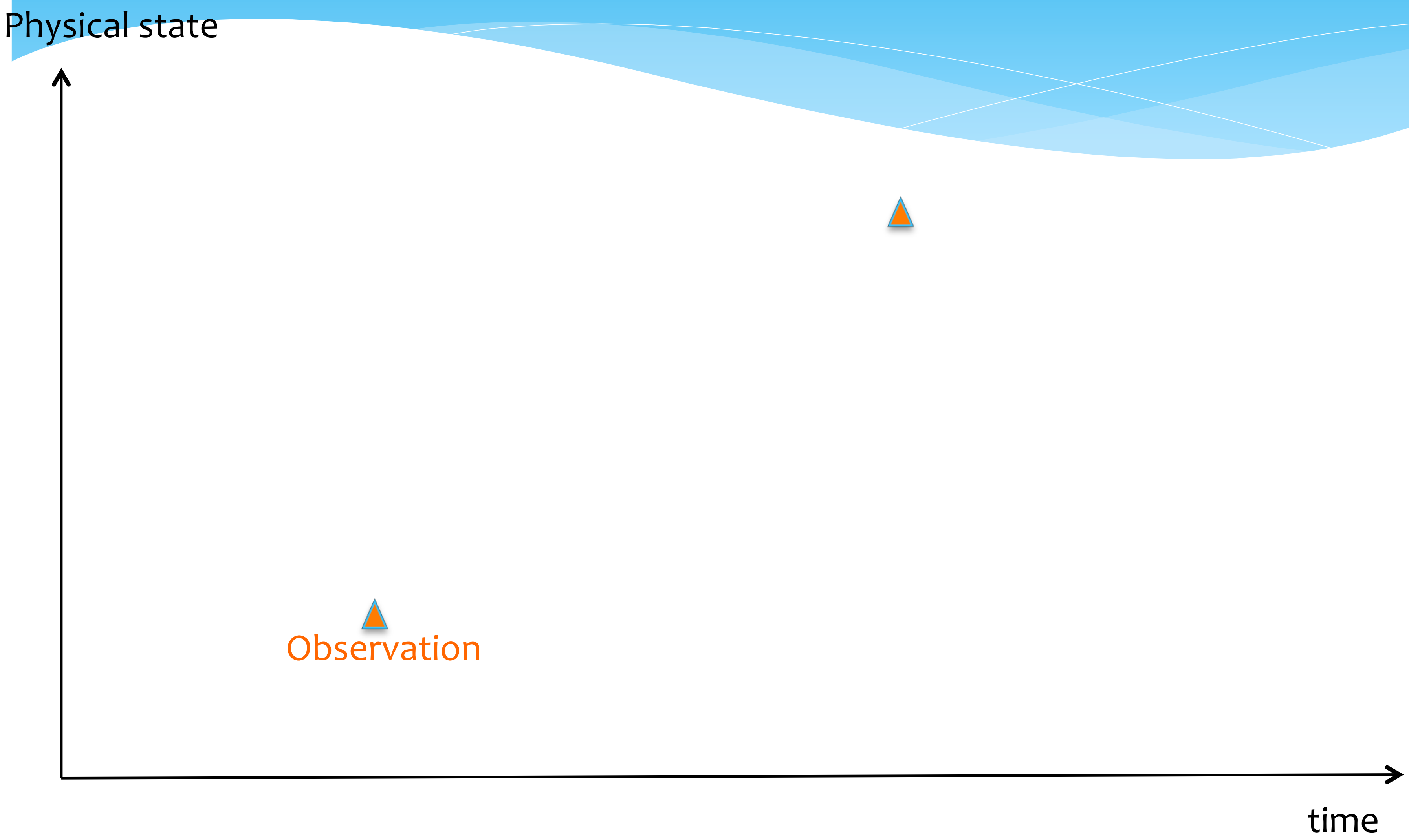
Forecast step:

Often nonlinear in practice

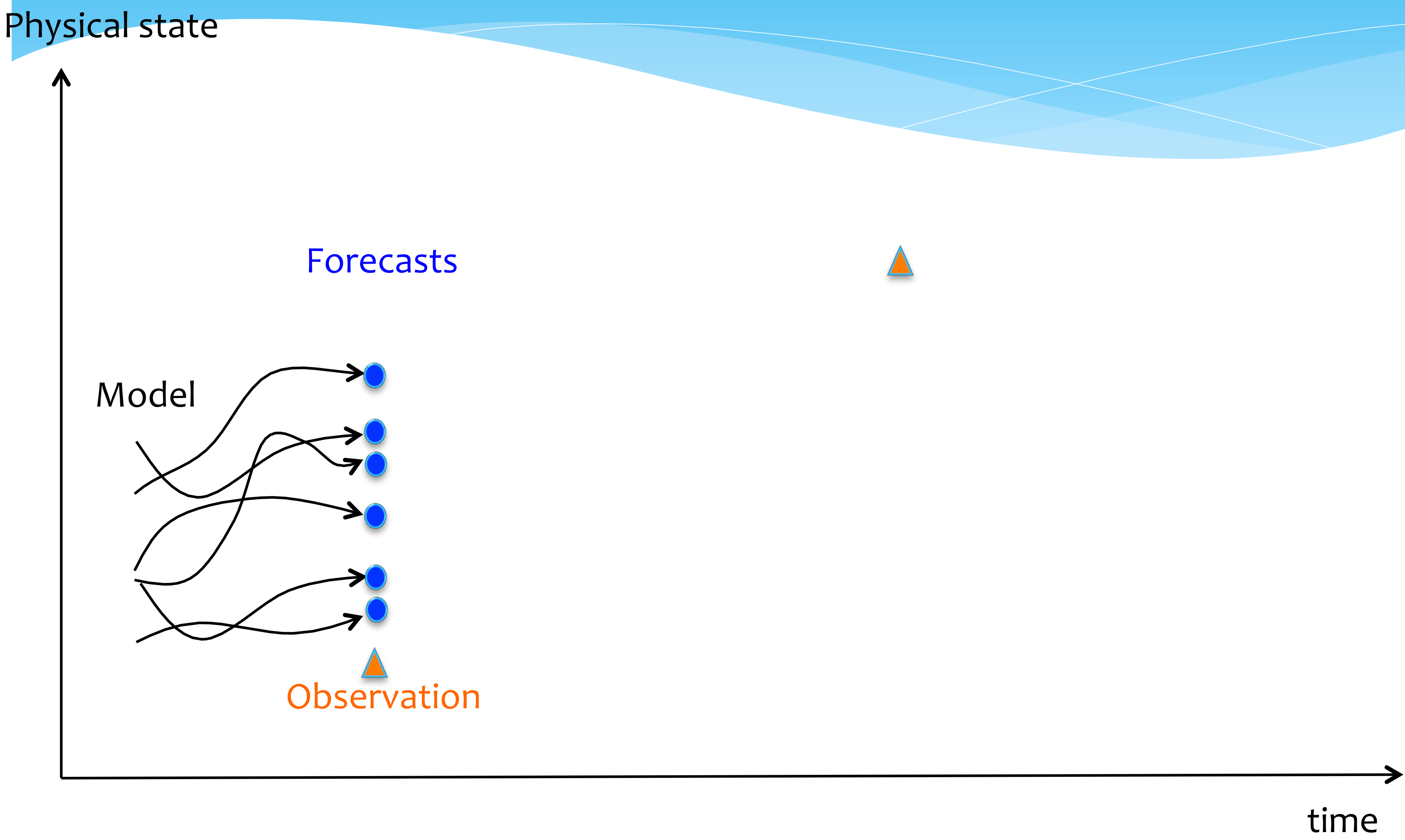
$$\begin{aligned} \mathbf{x}_{k+1}^f &= \mathbf{M}_{k,k+1}\mathbf{x}_k^a, \\ \mathbf{P}_{k+1}^f &= \mathbf{M}_{k,k+1}\mathbf{P}_k^a\mathbf{M}_{k,k+1}^T + \mathbf{Q}_k. \end{aligned}$$

Rarely that simple, and unknown

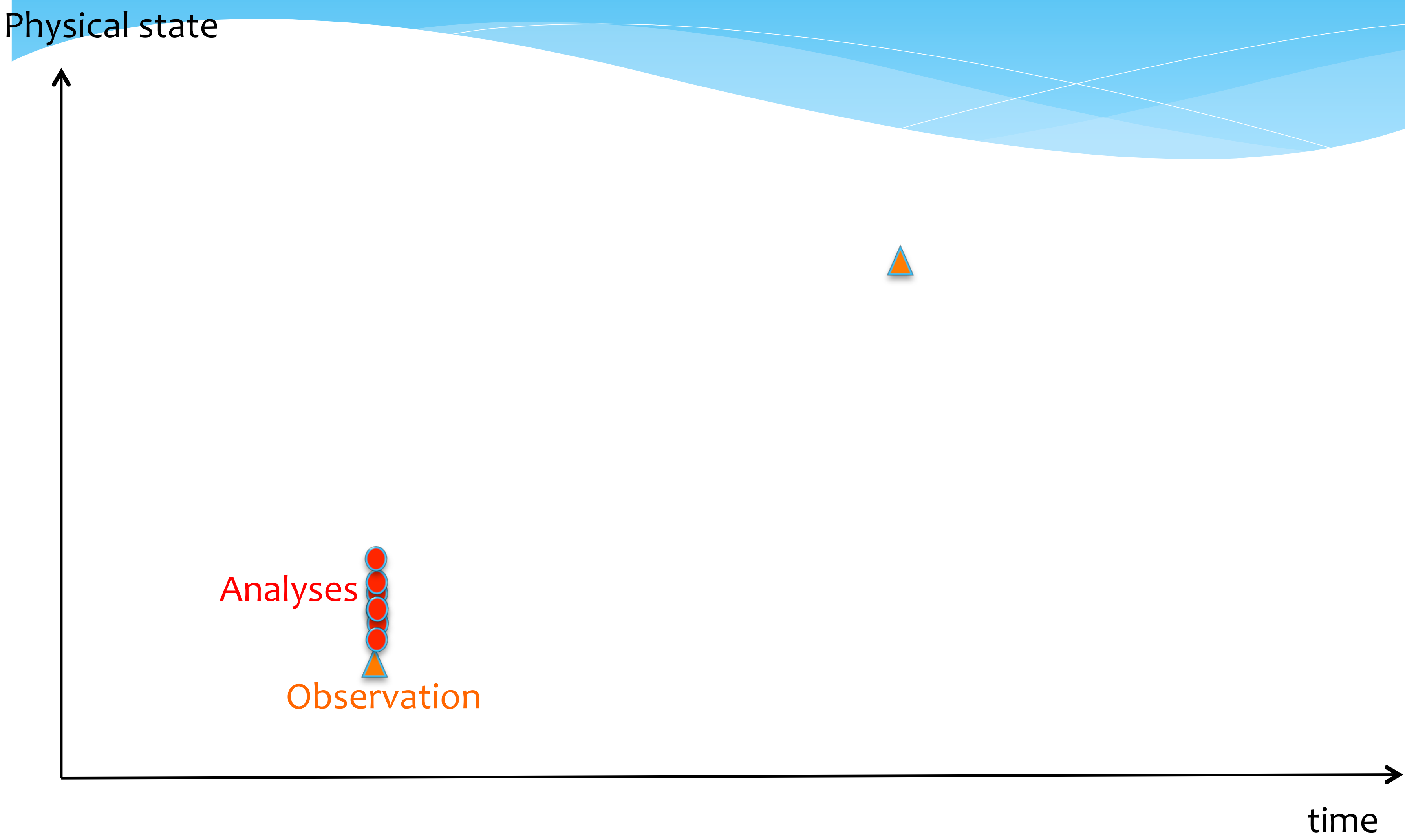
Ensemble Kalman filters



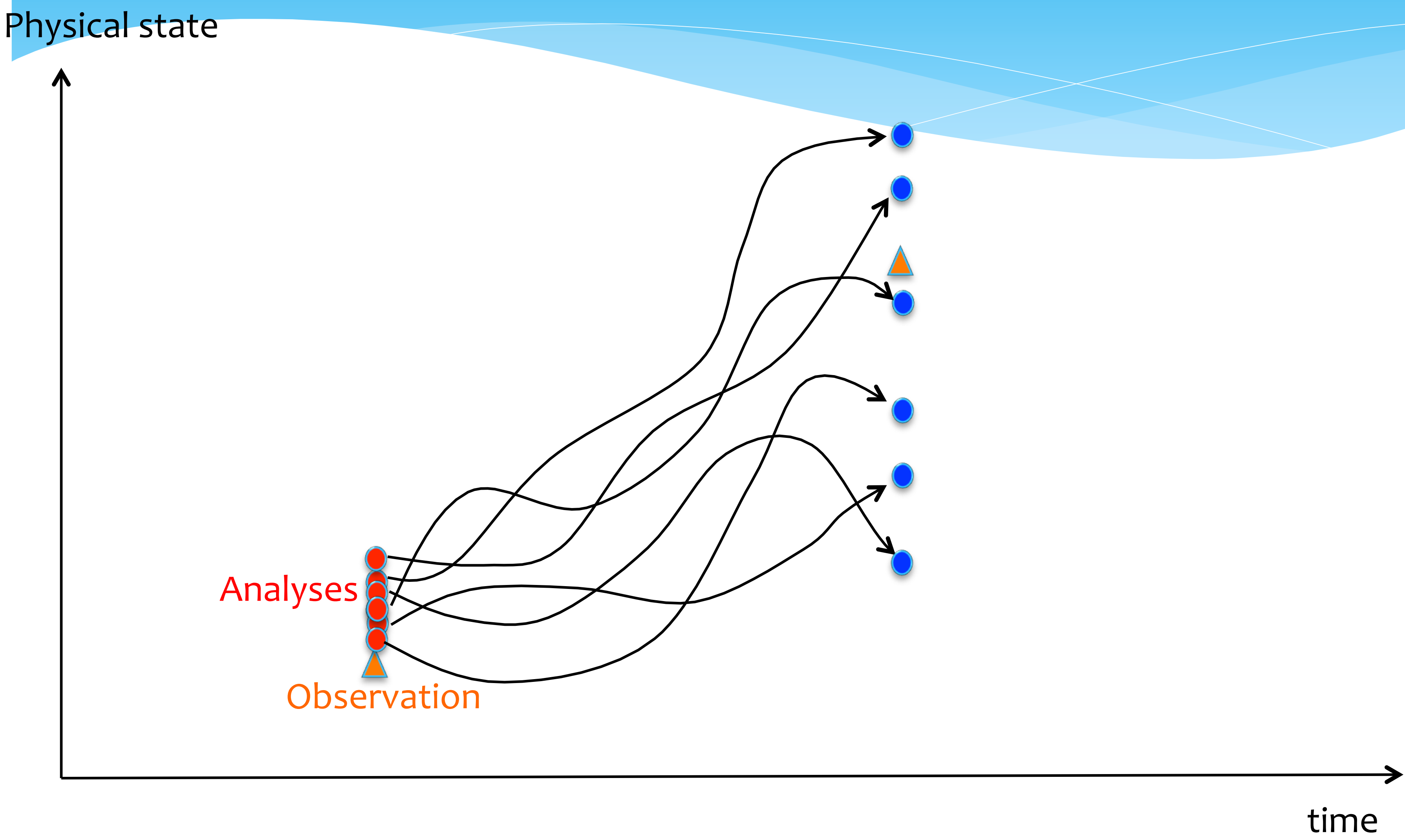
Ensemble Kalman filters



Ensemble Kalman filters

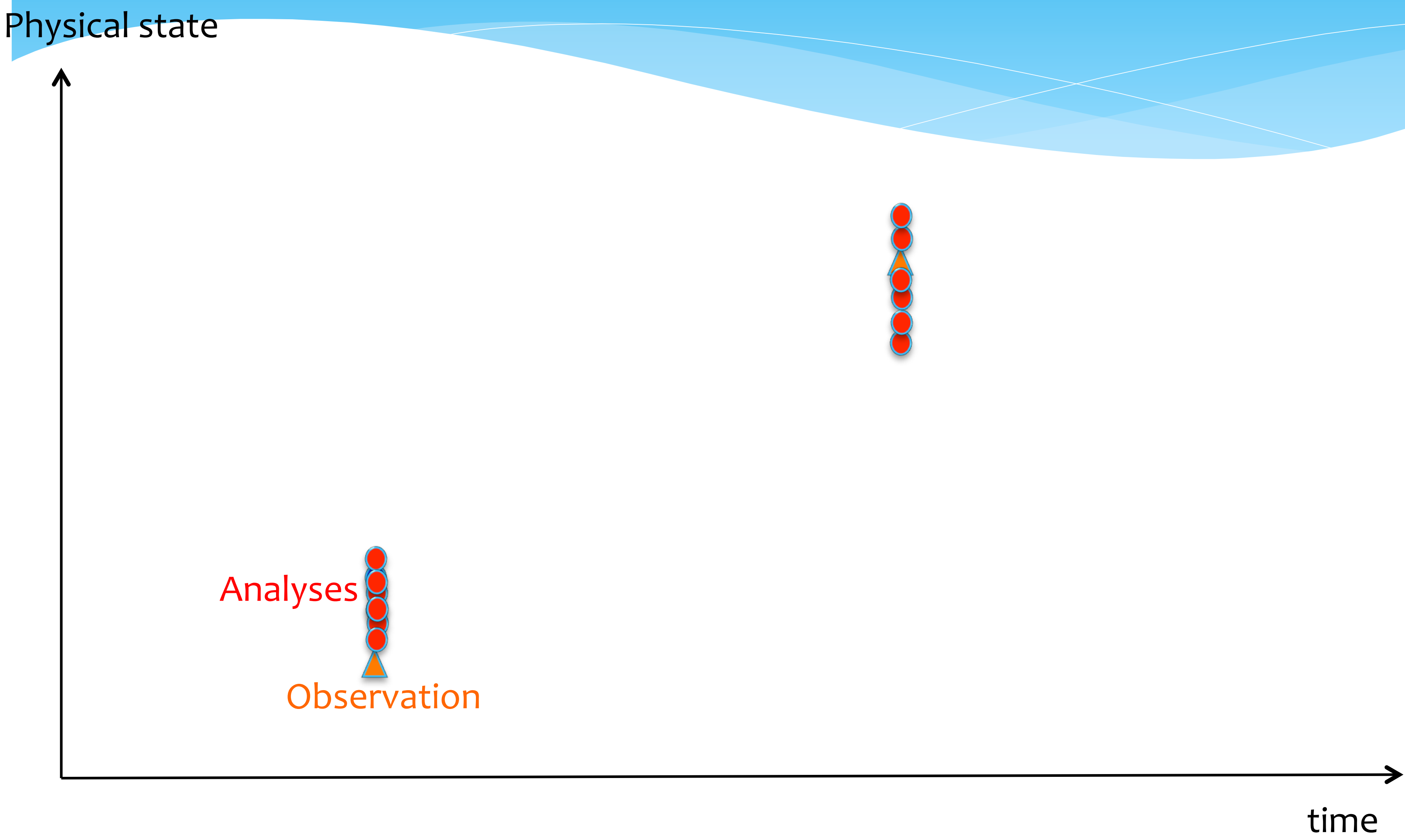


Ensemble Kalman filters

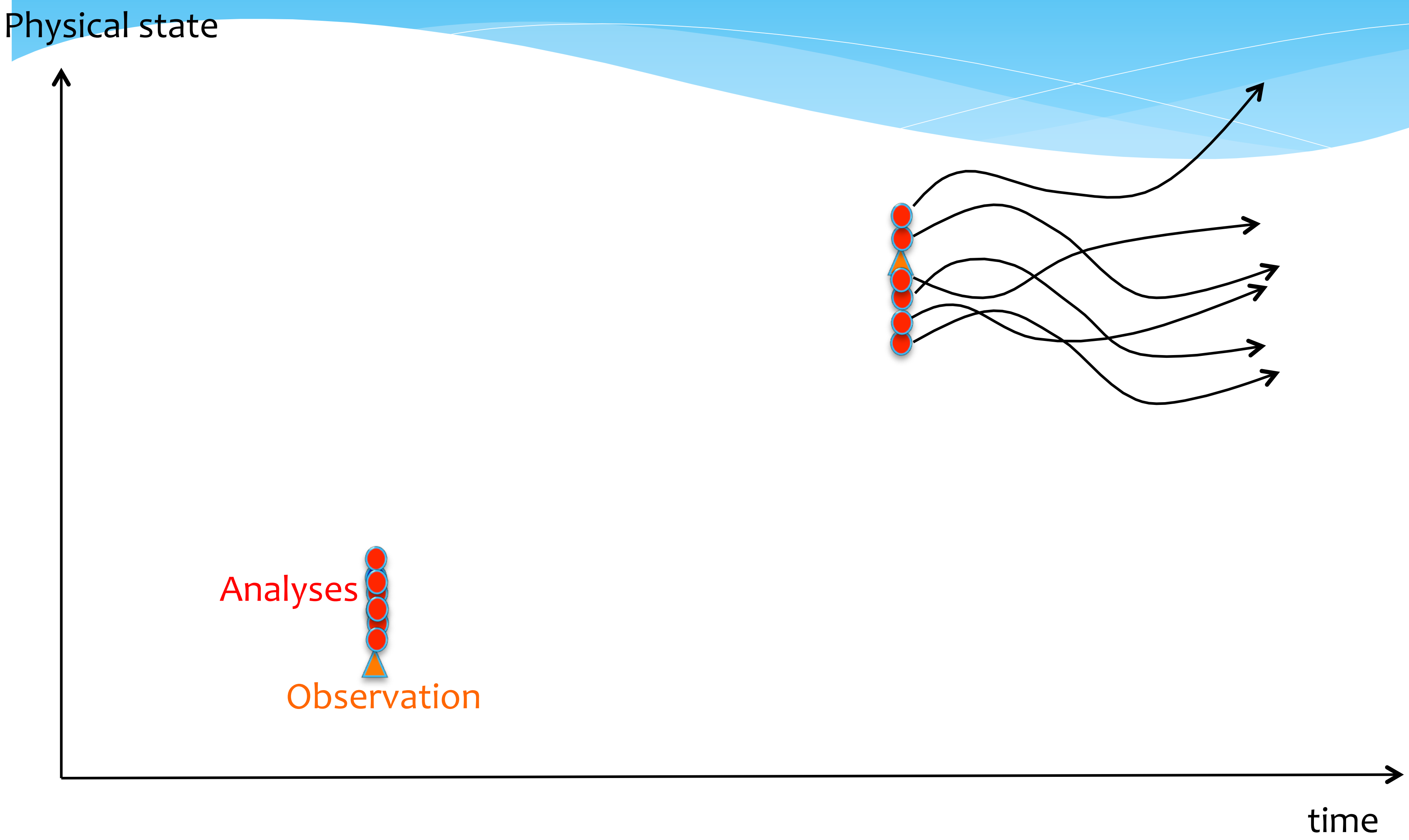


Ocean DA using EnKFs

Ensemble Kalman filters



Ensemble Kalman filters



Ensemble Kalman filters

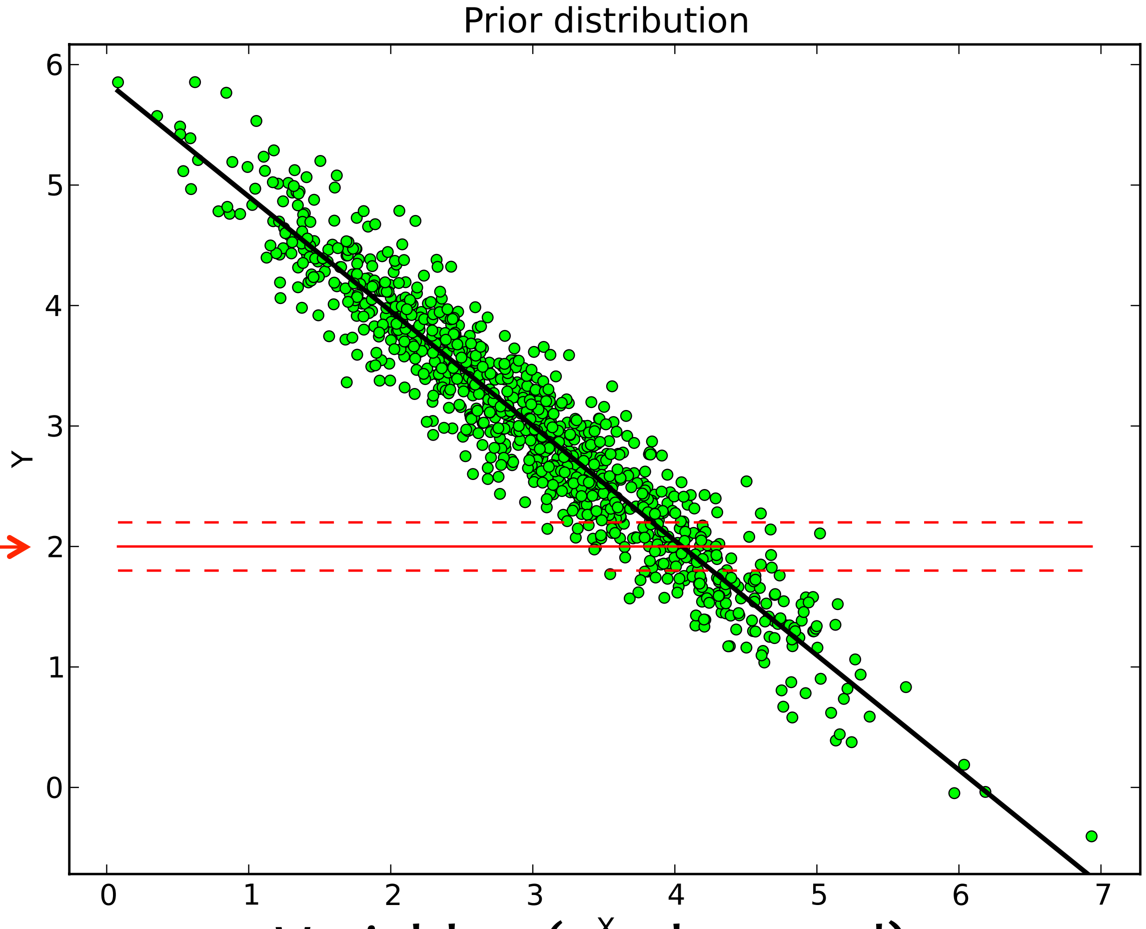
* In the forecast step, each member is advanced with the numerical model:

$$\mathbf{x}^f_{k+1,i} = M_{k,k+1}(\mathbf{x}^a_{k,i}) + \eta_{k,i}$$

Ensemble Kalman filters

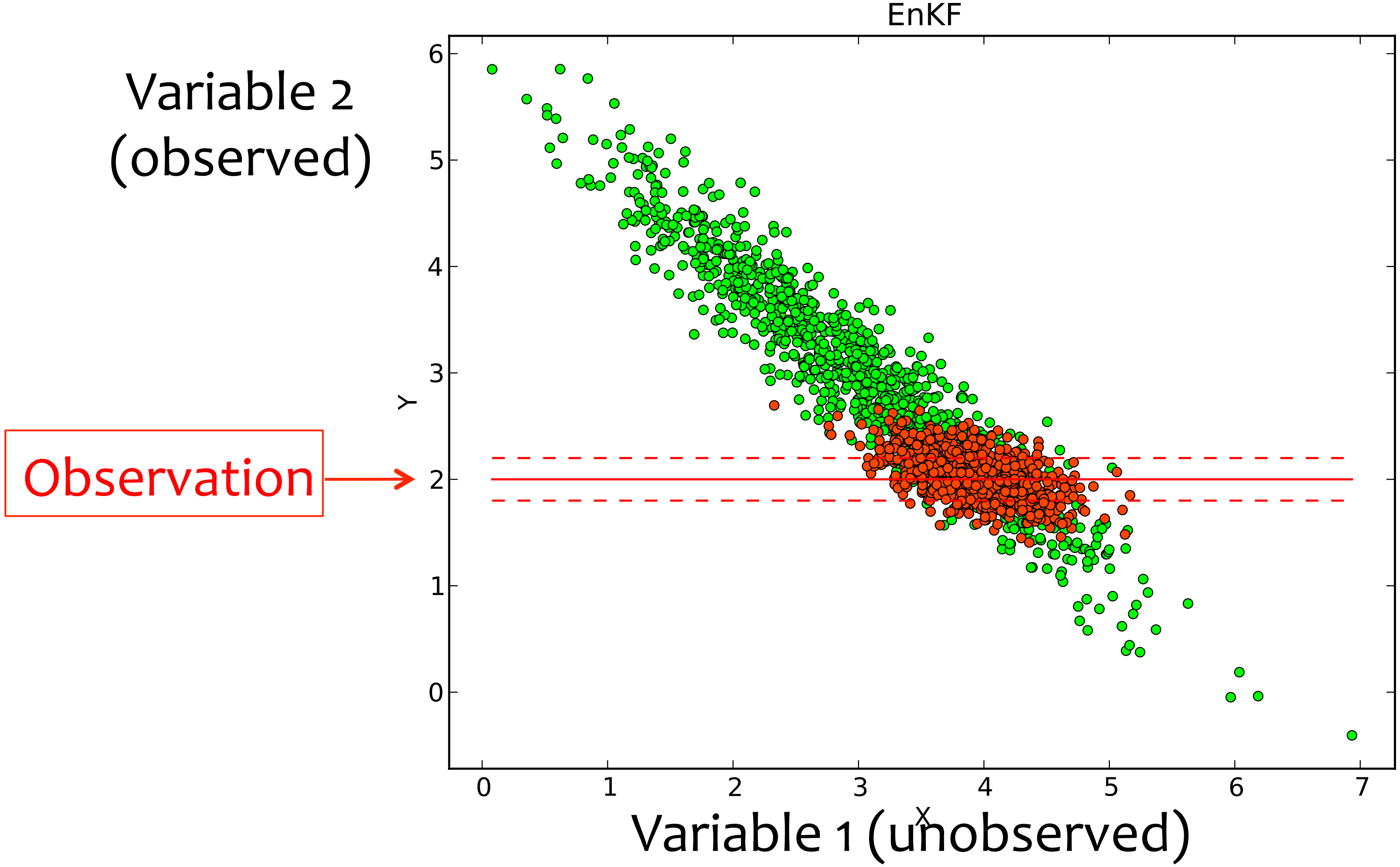
Variable 2 (observed)

Observation



Variable 1 (unobserved)

Ensemble Kalman filters



Ensemble Kalman filters

* At the analysis step, each member is corrected using observations.
* Different analysis schemes exist:
  * stochastic/deterministic,
  * algebra in observation/ensemble space,
  * Serial/batch processing of observations,
  * With/without adaptive scheme at some point,
  * etc

Ensemble Kalman filters

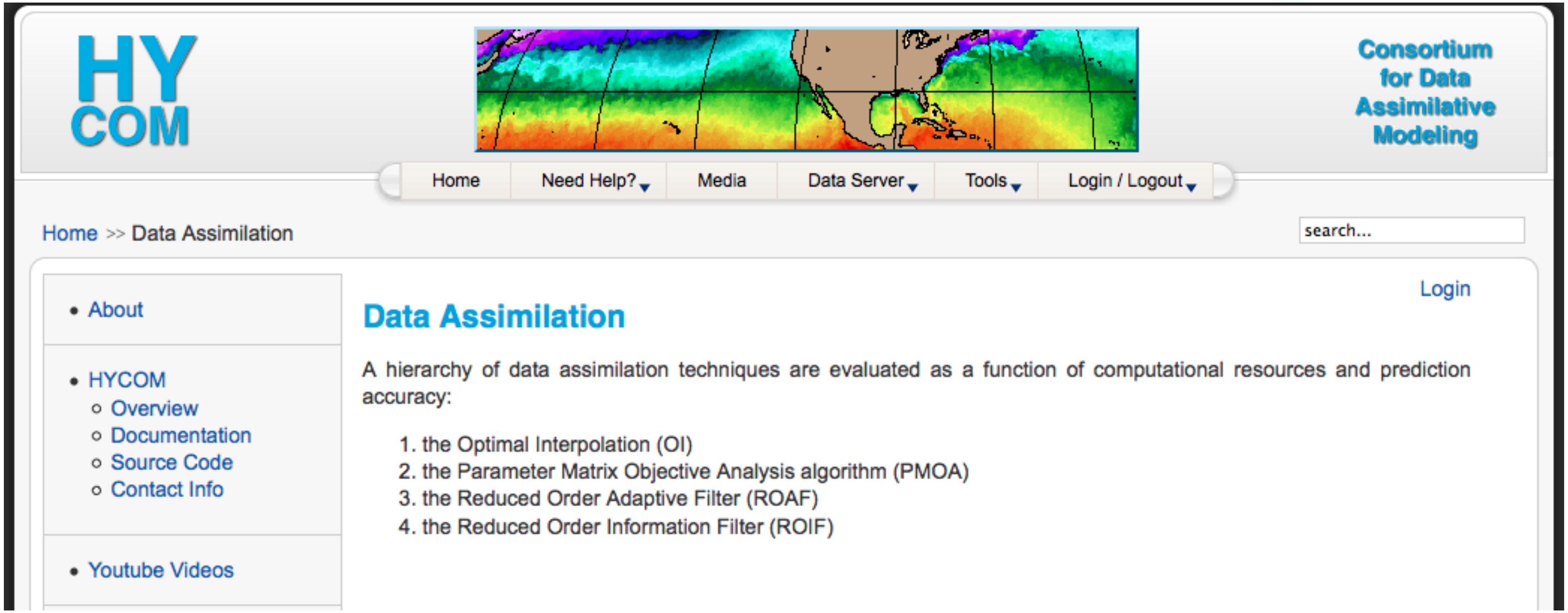
Deliverable 3.1

SANGOMA

Contents

SANGOMA European project, http://www.data-assimilation.net/)

Ensemble Kalman filters



http://hycom.org/

Ensemble Kalman filters

A simple view

* OI methods
  * Forecast of 1 (mean) state
  * Analysis using statistics from a fixed ensemble
* Stochastic EnKF
  * Correction of each state with perturbed observations
* Deterministic EnKFs
  * Correction of mean and anomalies without perturbing observations

Ensemble Kalman filters

* Ocean DA: $O(10^6 - 10^8)$ variables, $O(10^3 - 10^5)$ obs.
* Ensemble Kalman filters used in operational oceanic DA systems:
  * Ensemble OI (Mercator-Océan, France; Bureau of Meteorology, Australia; and others)
  * Deterministic EnKF (NERSC, Norway)

Ensemble Kalman filters

* Ensemble OI:
  * Only a mean state is propagated with the model;
  * The error modes are the same at any analysis step.
* - - - : no estimation of uncertainties;
* +++: computationally affordable, robust (no collapse), more "physically-based" than historical OI with analytical covariance functions.

Localization

* Localization aims at delimiting in space the impact an observation;
* Localization is necessary for several reasons:
  * To avoid long-range corrections due to spurious long-range correlations, themselves due to the small size of the ensemble;

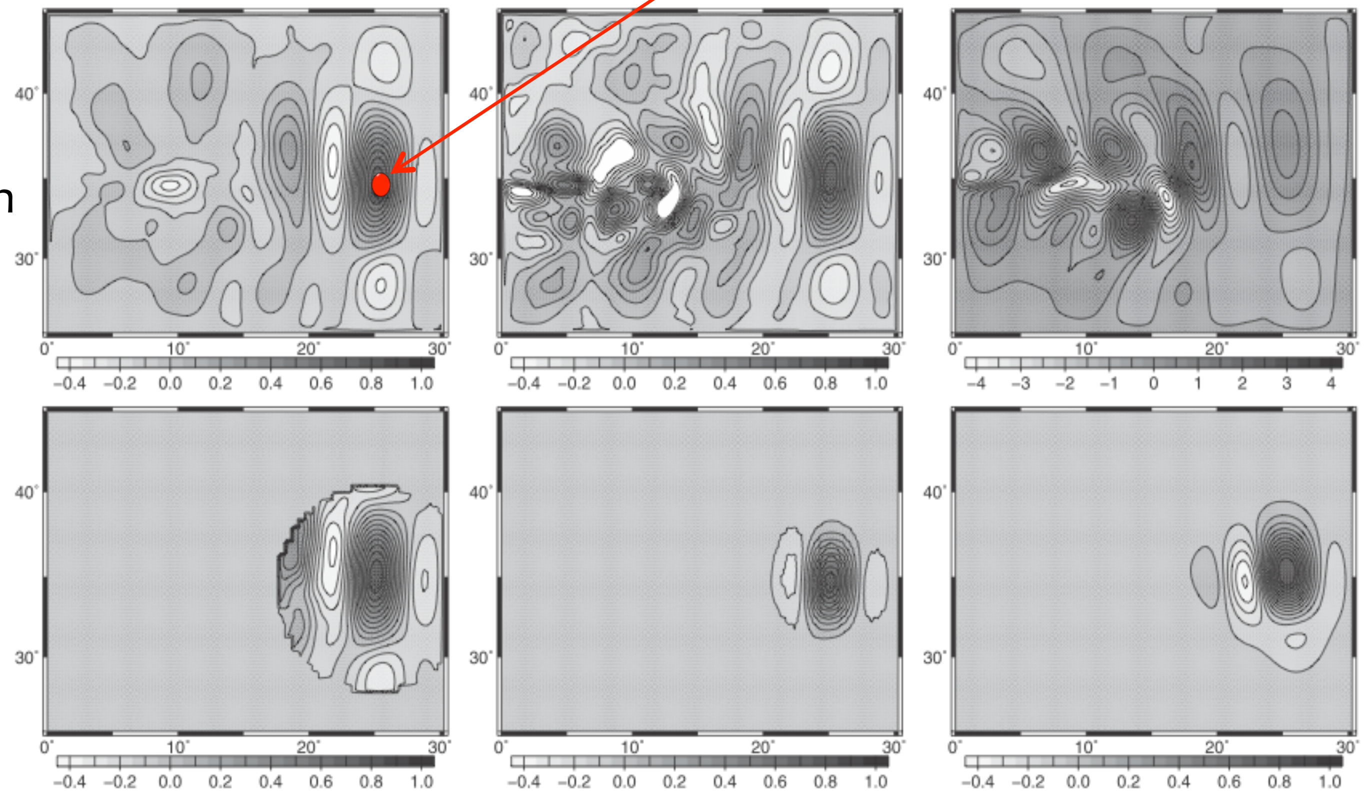
Localization

Increments in SSH due to an observation here

m=5000; no rank reduction | m=200; no rank reduction | m=5000; rank reduction r=20

Without localization

With localization



m=200; no rank reduction
Awkward localization

m=200; no rank reduction

m=5000; rank reduction r=20

(Brankart et al, 2011)

Localization

* Localization aims at delimiting in space the impact an observation;
* Localization is necessary for several reasons:
  * To avoid long-range corrections due to spurious long-range correlations, themselves due to the small size of the ensemble;
  * To artificially increase the rank of the covariance matrix and provide more degrees of freedom to the corrections.

Localization

* Why increasing the rank of the covariance matrix?
* Remember that in the ETKF, the correction on the mean is a linear combination of the anomalies (see Marc Bocquet's course, Eq. 5.34):

$$\mathbf{x}_i^a = \bar{\mathbf{x}}^f + \mathbf{X}^f \gamma_i$$

* There is only m (ensemble size, ~10-100) degrees of freedom to correct a vector of typical size > 10^6 with 10^3-10^5 observations!

Localization

- Localization aims at delimiting in space the impact of an observation;
- Localization is necessary for several reasons:
  - To avoid long-range corrections due to spurious long-range correlations, themselves due to the small size of the ensemble;
  - To artificially increase the rank of the covariance matrix and provide more degrees of freedom to the corrections;
  - To make computation possible in some cases.

Localization

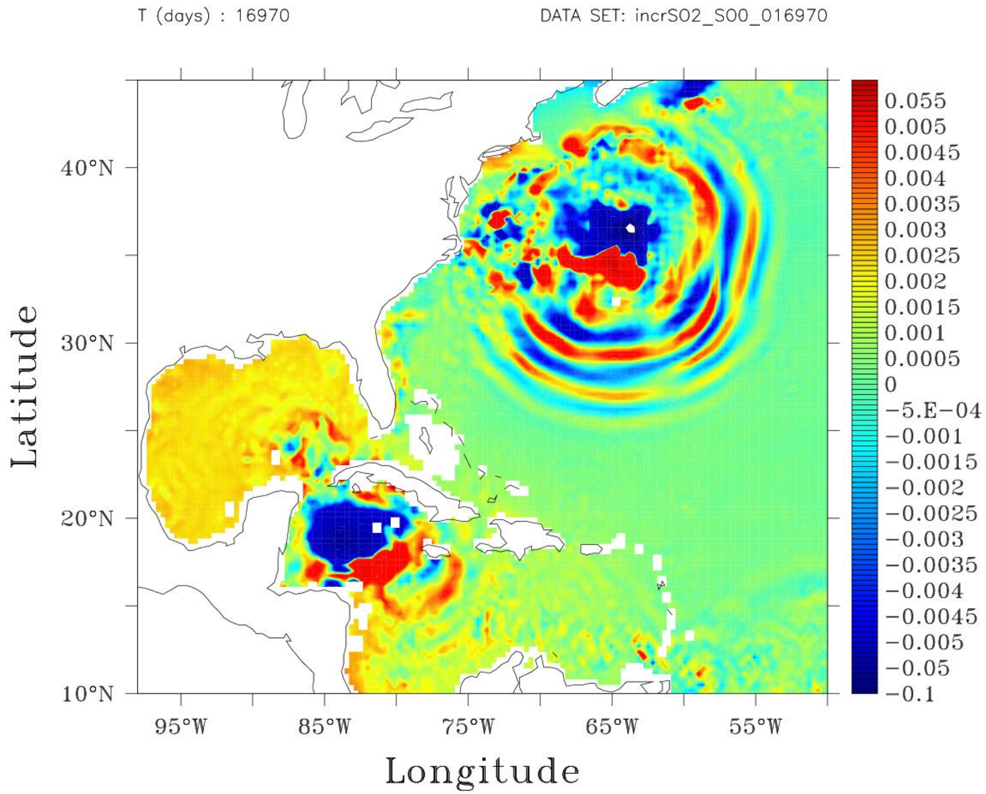
* The Kalman gain can be computed directly if the number of local observations (i.e., the size of *R*) is limited:

$$\mathbf{K}^{*} = \mathbf{P}^{\mathrm{f}}\mathbf{H}^{\mathrm{T}}\left(\mathbf{H}\mathbf{P}^{\mathrm{f}}\mathbf{H}^{\mathrm{T}} + \mathbf{R}\right)^{-1}$$

IAU

Model not involved during analysis: discontinuity, balance problems and shocks at restart possible.

Right: spurious wave generated by the assimilation of a single observation.



(Rozier et al, 2007)

IAU

* An empirical solution is Incremental Analysis Updating (IAU, Bloom et al, 1996)
* IAU consists in computing corrections at the analysis step, then re-running the ensemble over the forecast window, adding incrementally to each member its correction under the form of a forcing term.

IAU

Here, IAU is run from the middle of the previous forecast window to the middle of the next forecast window.

Continuity is guaranteed (perhaps at the expense of quality of the analysis).

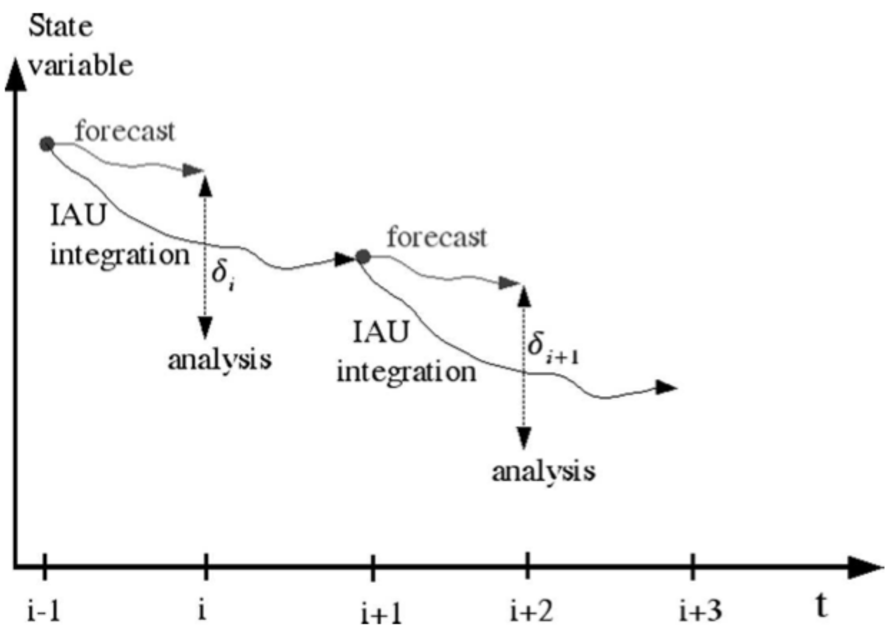


FIG. 1. IAU method from Bloom et al. (1996); δ represents the increment.

IAU

Figure: spatially averaged zonal velocity U in the Gulf Stream zone.
Black: free run
Red: EnOI
Green: EnOI with IAU

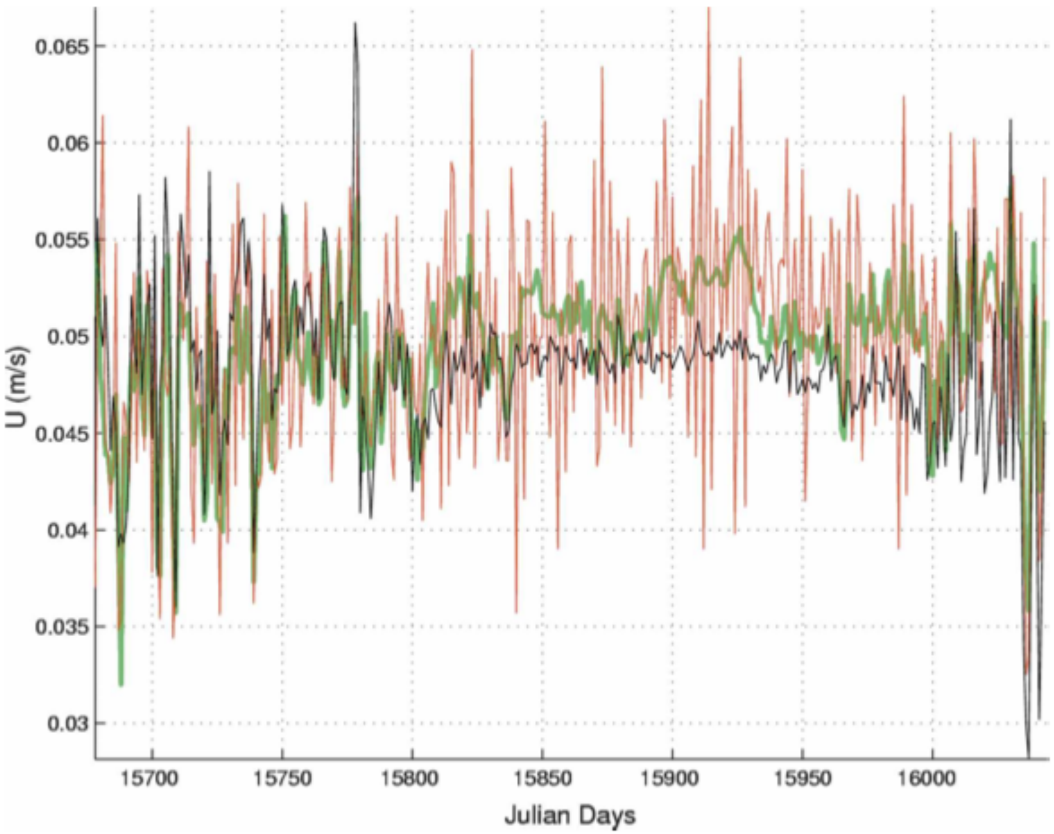


FIG. 12. Same as in Fig. 11, but at a 55-m depth (model depth level 5) from Julian day 15678 (4 Dec 1992) to 16038 (5 Dec 1993): black line represents FREE run, red line represents INT run, and green line represents IAU run.

(Ourmières et al, 2005)

Bogus

- Some quantities must be conserved. Example: mass.

$$\mathrm{div}\ \mathbf{u} = 0$$

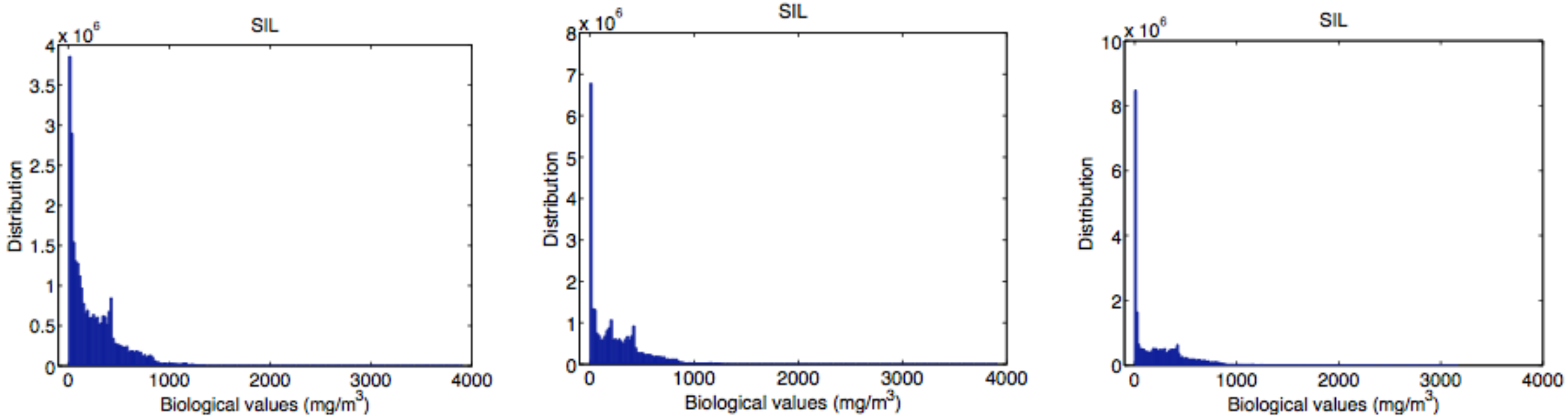
- Bogus: a fictitious observation of div **u**, equal to 0.

- Bogus can be used in regions where the assimilation makes things worse…

Gaussian anamorphosis

- Sometimes the distribution of some variables does not follow a Gaussian law:

Silicate



Distribution of silicate at 3 different dates (over a large oceanic domain)

(Simon et al, 2009)

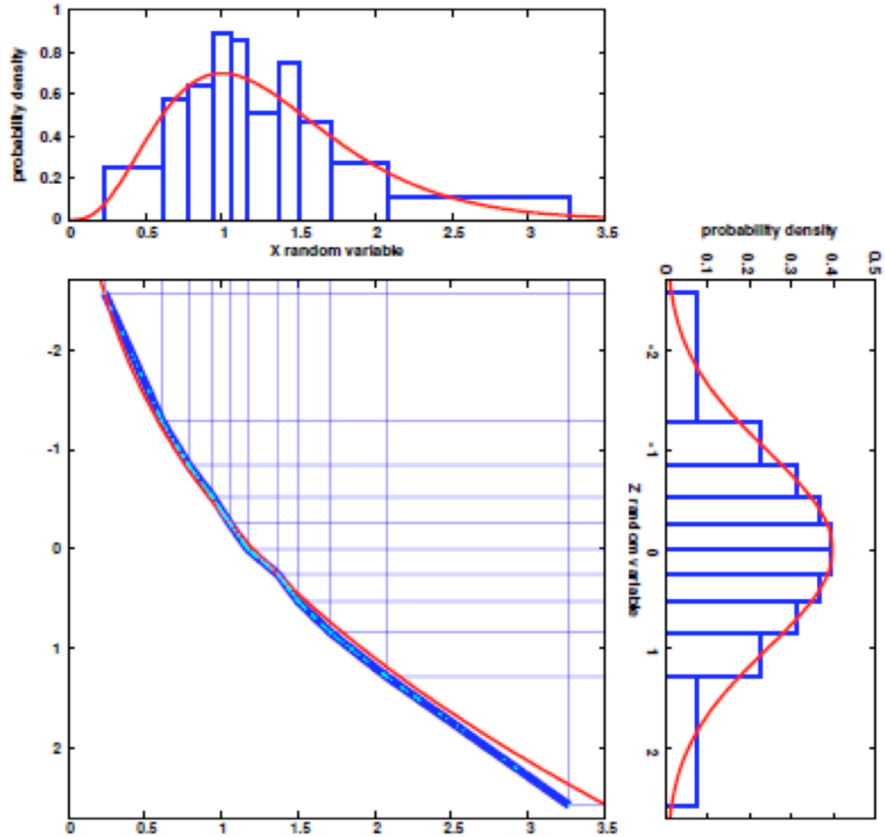
Gaussian anamorphosis

* Sometimes the distribution of some variables does not follow a Gaussian law;
* But the EnKFs work better with Gaussian variables;
* Gaussian anamorphosis: transformation of a distribution into a Gaussian distribution.

(Bertino et al, 2003)

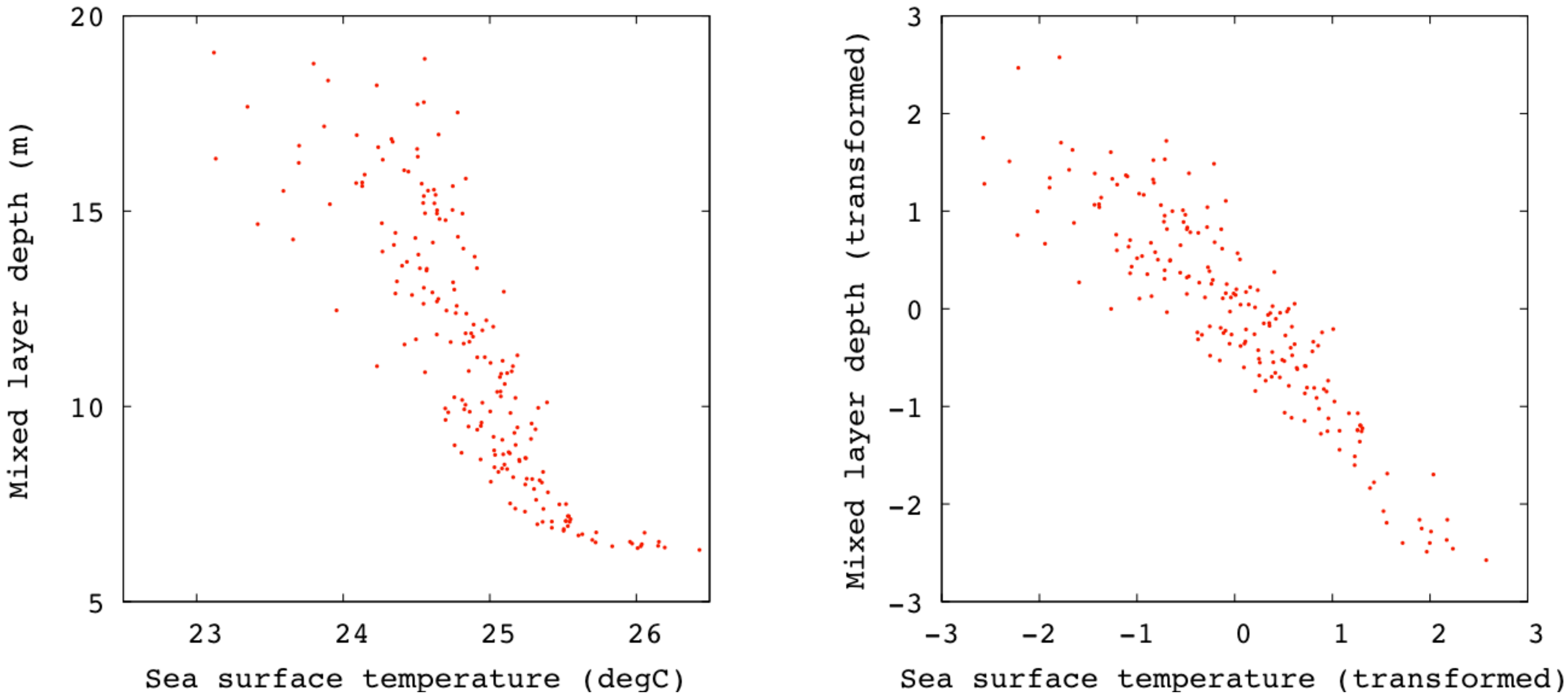
Gaussian anamorphosis

* The transformation can be analytical or empirical;
* On the opposite figure, the transformation is empirical;
* Such transformation can be performed on each variable individually.



(Béal et al, 2010)

Gaussian anamorphosis



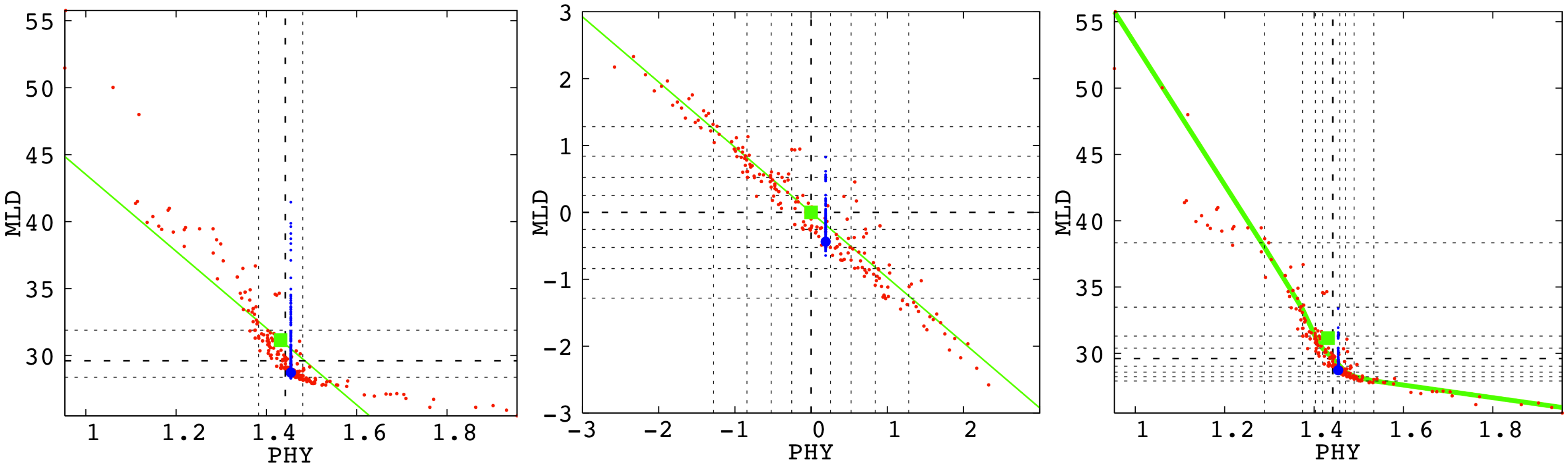
Here, the anamorphosis tends to "Gaussianize" the bivariate distribution.

(Brankart et al, 2012)

Gaussian anamorphosis

* After transformation, the EnKF analysis is performed;
* Then, the physical variables are retrieved by the inverse transformation.

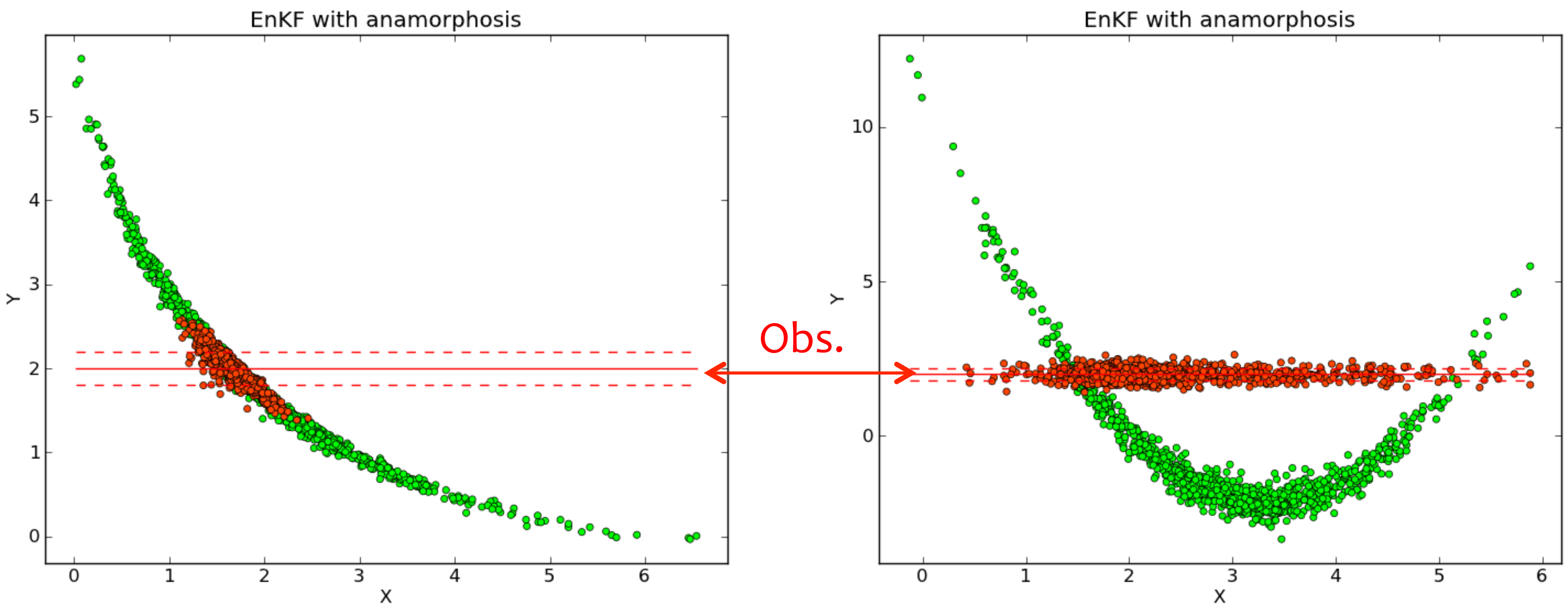
Gaussian anamorphosis



Obs. update at BATS station (65∘W-32∘N) using a perfect PHY observation. Prior ensemble (red), mean (green square), linear regression line (thin green line), truth (big blue dot), posterior ensemble (blue dots). Left: EnKF analysis; Middle: analysis in the transformed state space; Right: Anamorphosis-EnKF posterior. The thick green line on the right is the transformation of the thin green line on the middle.

(Béal et al, 2010)

Gaussian anamorphosis



Gaussian anamorphosis works well with weakly non Gaussian variables…

(Metref et al, 2014)

About the observation error covariance matrix

$$\mathbf{P}^f = \mathbf{S}^f {\mathbf{S}^f}^T$$

* The EnKF correction is either calculated with (using a serial processing of observations)

$$\delta\mathbf{x} = \mathbf{S}^f(\mathbf{HS}^f)^T\left[(\mathbf{HS}^f)(\mathbf{HS}^f)^T + \mathbf{R}\right]^{-1}(\mathbf{y} - \mathbf{Hx}^f),$$

* Or, with $\Gamma = (\mathbf{HS}^f)^T\mathbf{R}^{-1}(\mathbf{HS}^f)$

$$\delta\mathbf{x} = \mathbf{S}^f\left[\mathbf{I} + \Gamma\right]^{-1}(\mathbf{HS}^f)^T\mathbf{R}^{-1}(\mathbf{y} - \mathbf{Hx}^f).$$

About the observation error covariance matrix

* For simplification, all ocean DA systems consider the observation error covariance matrix diagonal.
* To minimize the impact of the neglected correlations, it is common to inflate the variances (in the Norwegian operational system, they are multiplied by 2 for the update of the anomalies).
* On the other hand, many efforts are dedicated to the construction of the state error covariance matrix.

Ocean DA using variational methods

* Variational methods
* Incremental 4DVar
* Parameterization of the covariance matrix
* On Ensemble/variational methods

Variational methods

* Problem posed as the minimization of a cost function to find the best compromise between a prior knowledge x^b and observations y:

$$J(x) = \underbrace{\boxed{\frac{1}{2}\|x - x^b\|_b^2}}_{J_b} + \underbrace{\boxed{\frac{1}{2}\|H(x) - y\|_o^2}}_{J_o}$$

* With respect to a control vector x to choose carefully (very often: initial condition)

Variational methods

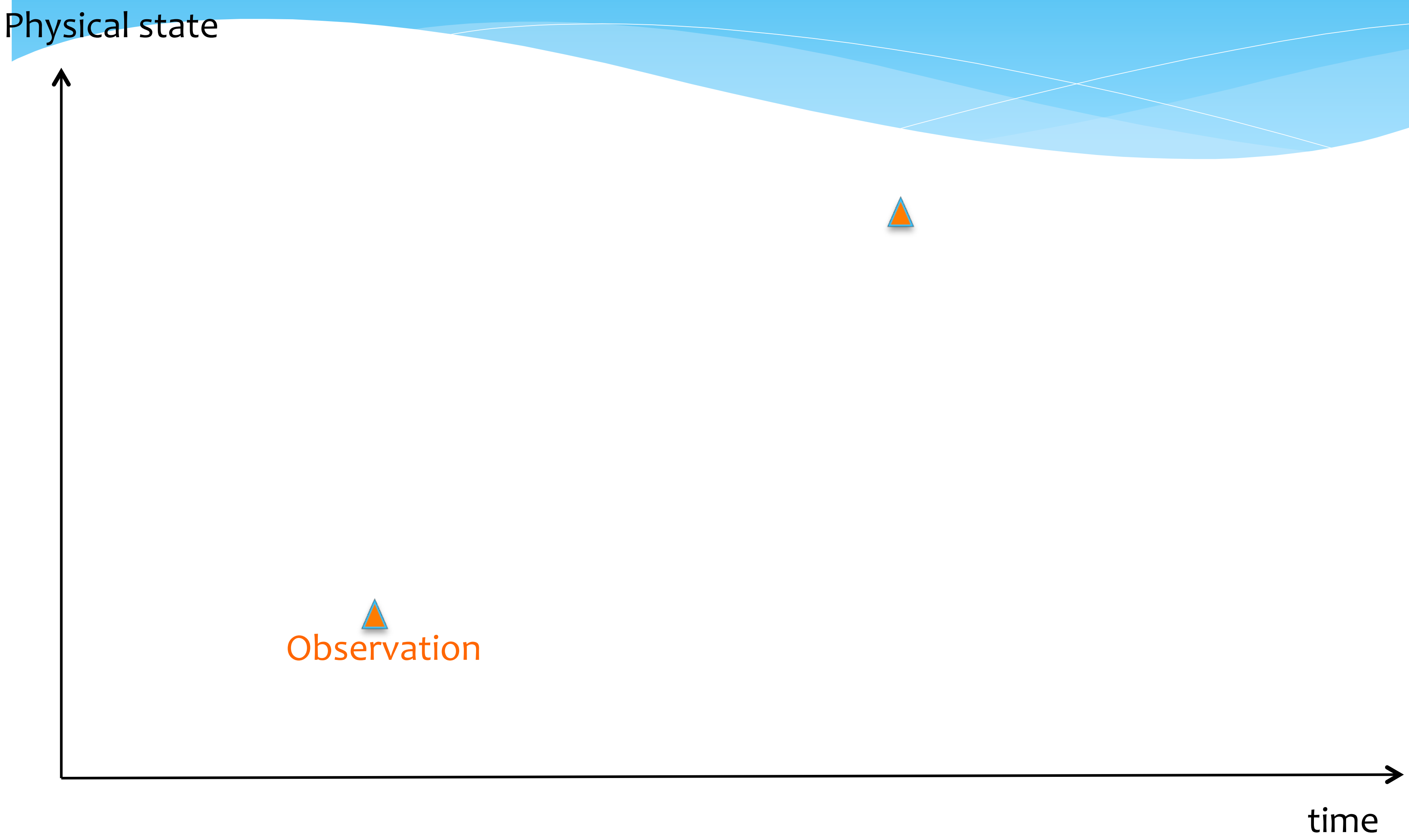
- 3DVar and 4DVar: the cost functions are quadratic.

$$J_{3D}(x_0) = \frac{1}{2}(x_0 - x^b)^T \mathbf{B}^{-1}(x_0 - x^b) + \frac{1}{2}(H(x_0) - y_0)^T \mathbf{R}^{-1}(H(x_0) - y_0)$$

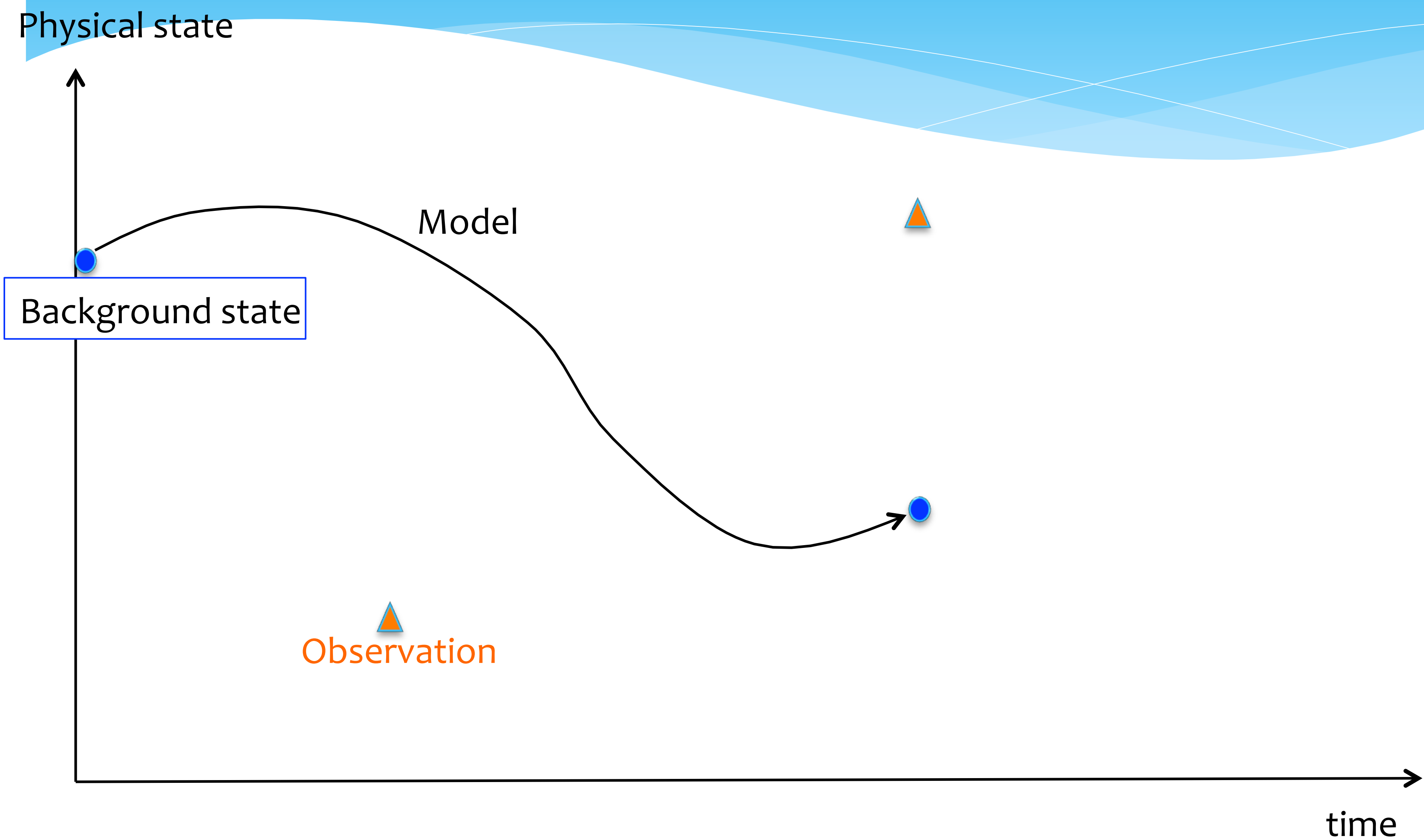
$$J_{4D}(x_0) = \frac{1}{2}(x_0 - x^b)^T \mathbf{B}^{-1}(x_0 - x^b) + \frac{1}{2}\sum_{i=0}^{N}(H(M_{0\to i}(x_0)) - y_i)^T \mathbf{R}^{-1}(H(M_{0\to i}(x_0)) - y_i)$$

- Efficient minimisation algorithms are iterative and require the gradient $\nabla J(x_0)$
- Adjoint methods are (by far) the cheapest ways to compute the gradient at each iteration.
- The adjoint model is often 2-4 times more expensive than the direct model.

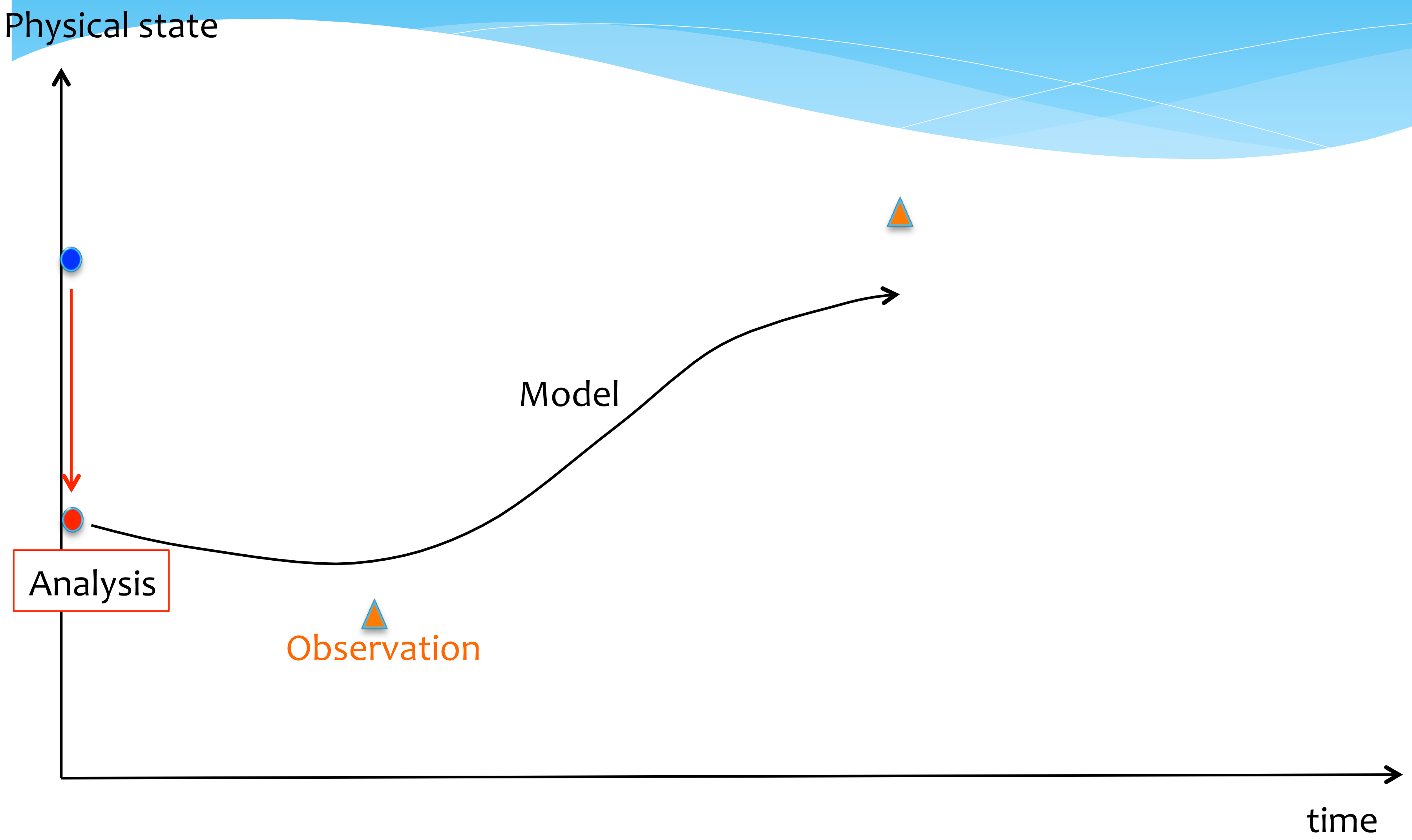
Variational methods



Variational methods



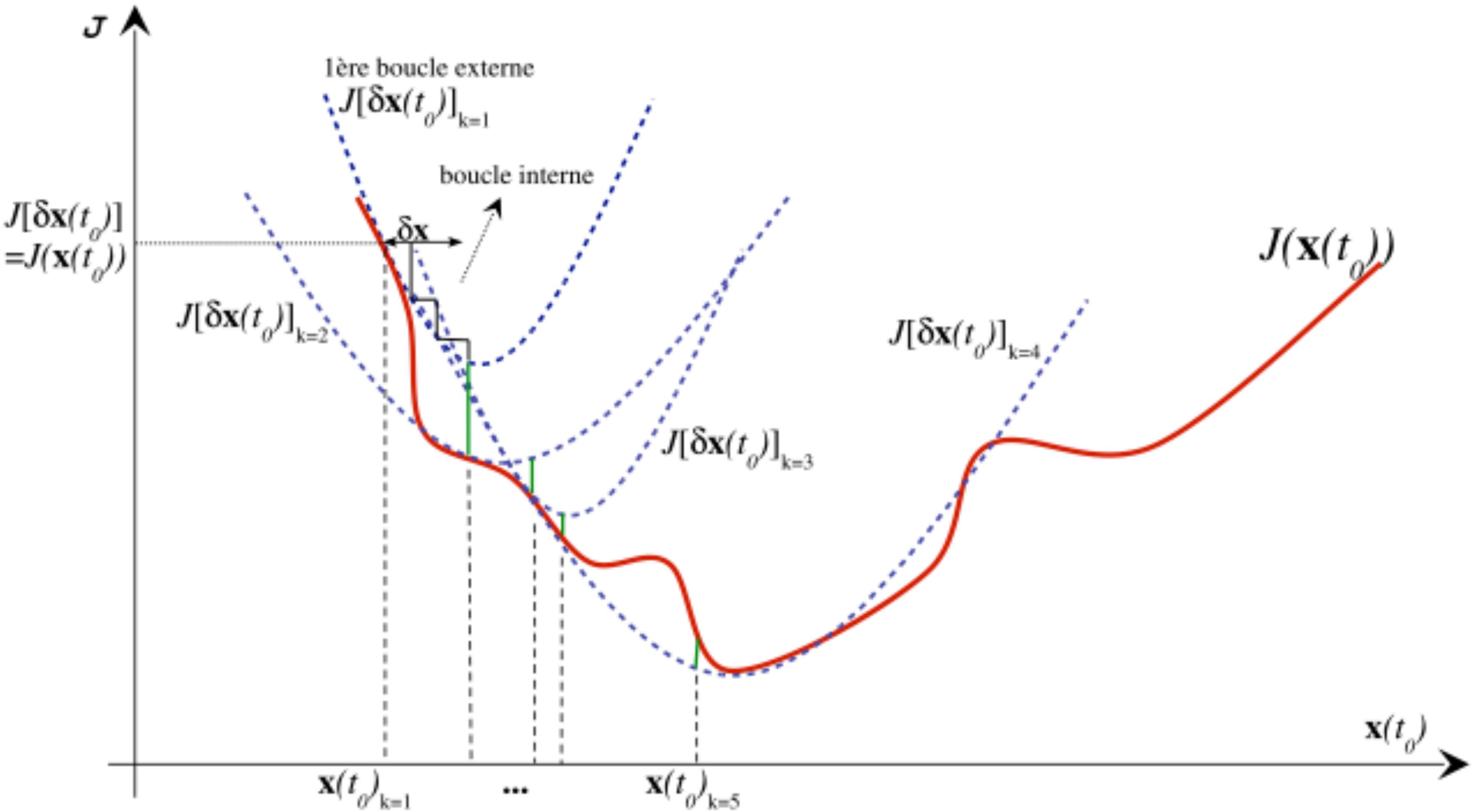
Variational methods



Incremental 4DVar

* When the model is non-linear, the cost function can be non-convex.
* Incremental 4D-Var splits the minimisation problem into a series of minimisations of quadratic (convex) cost functions.
* This leads to define outer and inner loops in the minimisation process.

Incremental 4DVar



Parameterization of the covariance matrix

* As with the EnKF, the full covariance matrix cannot be built and stored.

Parameterization of the covariance matrix

* A reduced-rank approach can be considered.
* The 4DVar increment is searched as a linear combination of a fixed set of error modes:

$$\delta \mathbf{x}_0 = \sum_{i=1}^{r} w_i \mathbf{L}_{\{i\}} = \mathbf{L}\mathbf{w}$$

* Minimization is carried out on **w**, a vector of size r.

(Robert et al, 2005)

Ocean DA using Var.

Parameterization of the covariance matrix

Experiment with a Tropical Atlantic model and 1 observation of T. Figure shows the increment in T.

Maximal correction is
0.94 on top
0.06 on bottom

The observation is here

Vertical section

Full 4DVar (diagonal B)

Reduced rank 4DVar

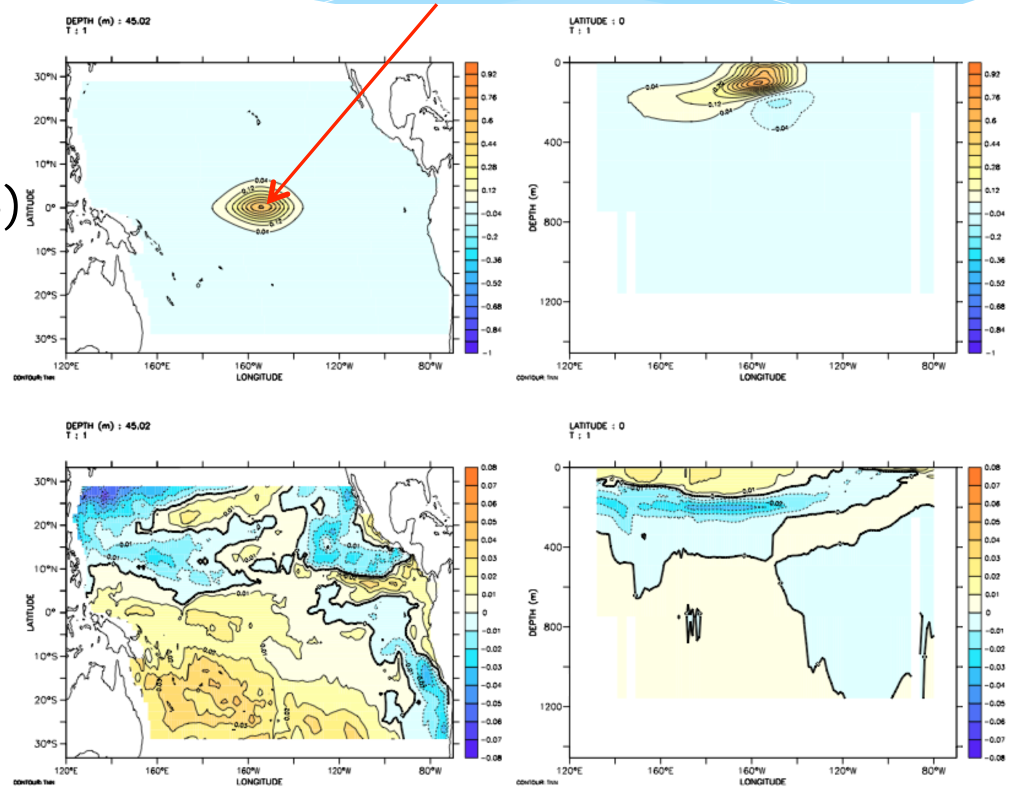


Fig. 4. Temperature component of the optimal increment δx_0 for single observation experiments. Left: horizontal structure at $z = -45$ m; right: vertical section along the equator. Top: full-space 4D-Var; bottom: reduced-space 4D-Var.

(Robert et al, 2005)

Parameterization of the covariance matrix

* Modelling of the covariance matrix with a suite of operators:

$$B = KD^{1/2}C^{1/2}(C^{1/2})^T D^{1/2}K^T$$

with

* K: balance operator
* D: variances (diagonal)
* C: correlations (block diagonal), built with a diffusion operator

(Weaver et al, 2005)

Parameterization of the covariance matrix

* The balance operator is introduced to form uncorrelated variables from the physical variables:

$$(T, S, SSH, U, V) \xrightarrow{K^{-1}} (T, S_U, SSH_U, U_U, V_U)$$

* The uncorrelated variables are then used in the control vector.
* The uncorrelated (unbalanced) variables are formed by removing their parts that are balanced by the others.

(Weaver et al, 2005)

Ocean DA using Var.

Parameterization of the covariance matrix

A single obs of T, located at 160W, 0N, 100 m depth. 10-day 4DVar increments on SSH, without (left) and with (right) the balance operator.

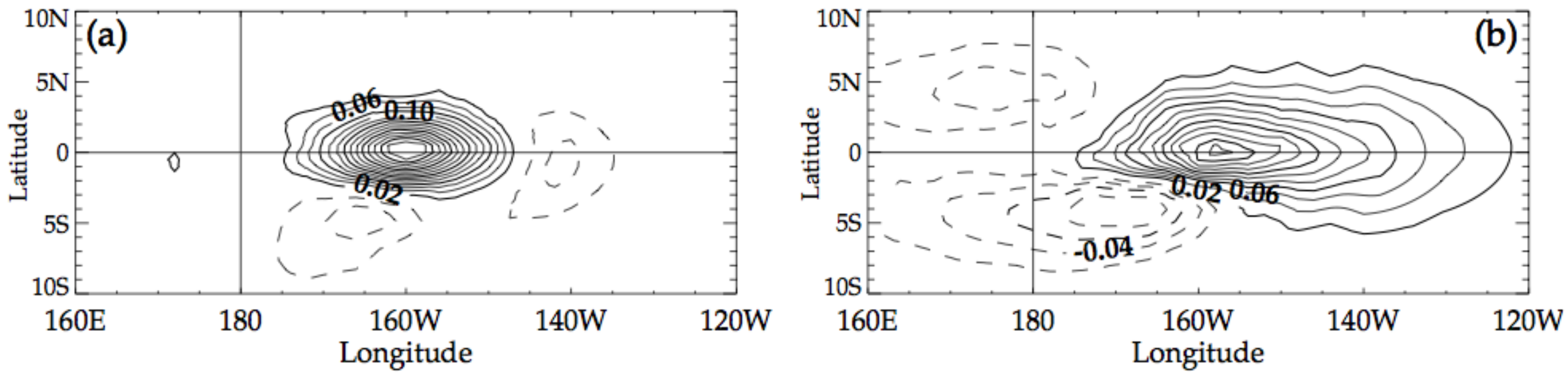


Figure 4. Horizontal section of the SSH analysis increments generated by the 4D-Var assimilation of a single-temperature observation (positive innovation) located ten days into an assimilation window at the same geographical location as in the example in Fig. 2. The increments are displayed on day 10 for a 4D-Var experiment (a) without and (b) with the balance operator activated. The fields have been multiplied by a factor 100 and the same contour interval has been used here as in Fig. 2(e). Solid (dashed) contours indicate positive (negative) values.

(Weaver et al, 2005)

Future challenges

* Consolidating Ensemble, 4DVar and/or Ensemble/variational methods?
* Modeling the model error with stochastic parameterizations
* Assimilation of images
* SWOT

Consolidating Ensemble, 4DVar and/ or Ensemble/variational methods?

* Ensemble methods are difficult to implement for high-resolution, realistic oceanographic problems:
  * problem size w.r.t. CPU capacities;
  * Model error;
* Adjoint-based DA methods are difficult to implement too:
  * Missing workforce for development and maintenance
  * Adjoint model hard to simplify (cost due to resolution)
  * Limited number of obs: strong nonlinearities develop

Modeling the model error with stochastic parameterizations

* Stochastic parameterizations emerge as an essential tool for representing uncertainties in models, hence for DA.
* But the set-up can be challenging.

Modeling the model error with stochastic parameterizations

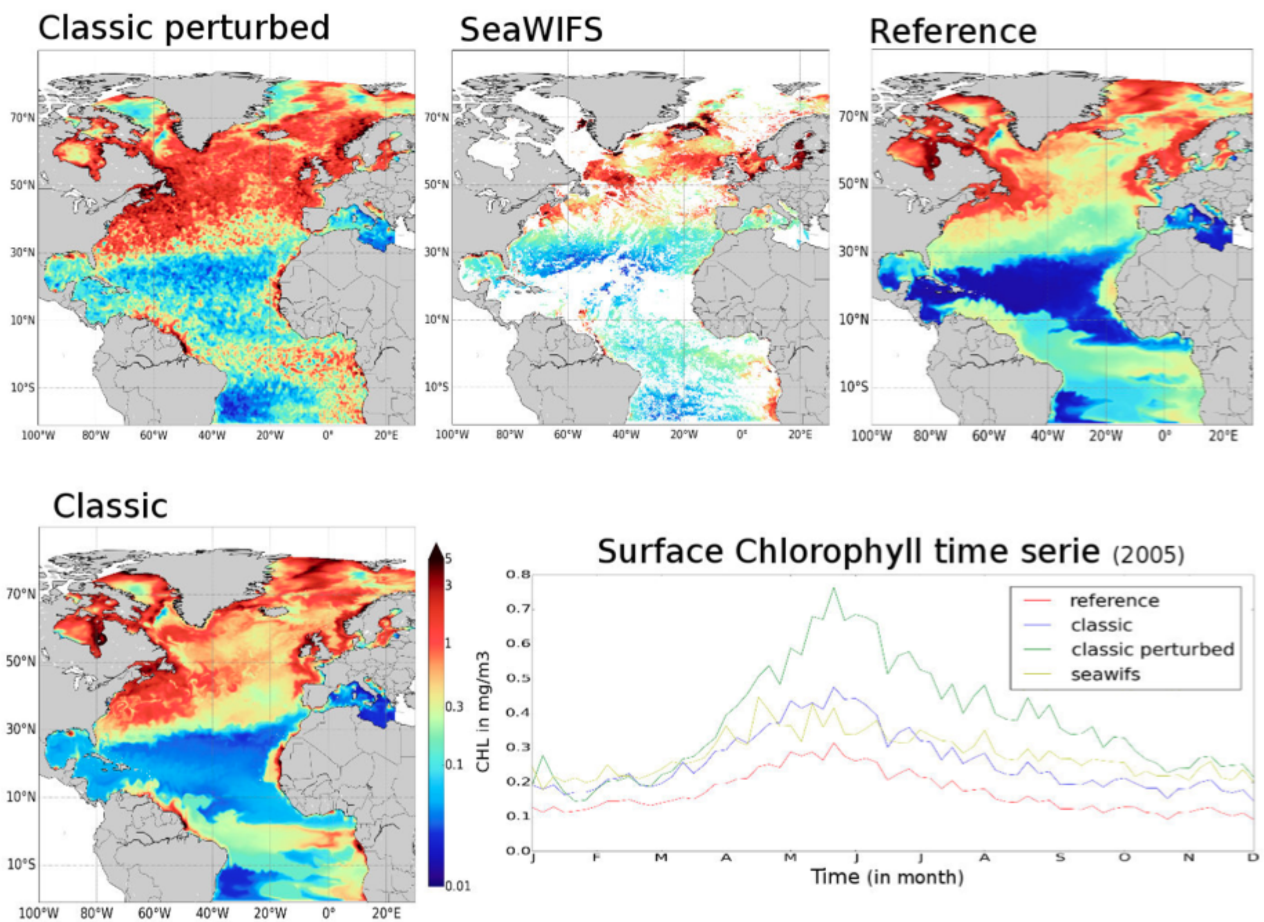
Example:
Ensemble of simulations of the biogeochemical cycle in the North Atlantic
Perturbation of 7 parameters (phyto growing rate, grazing rate, etc) around the values prescribed originally in the "deterministic" model.

$$\frac{\partial c_i}{\partial t} + \boldsymbol{u}.\nabla^h c_i + w\frac{\partial c_i}{\partial z} = \frac{\partial}{\partial z}\left(\tilde{\lambda}\frac{\partial c_i}{\partial z}\right) + S(c_i, c_j, \alpha_n)$$

(Garnier et al, 2016)

Modeling the model error with stochastic parameterizations

A correct representation of phytoplankton requires a full retuning of the reference parameters.



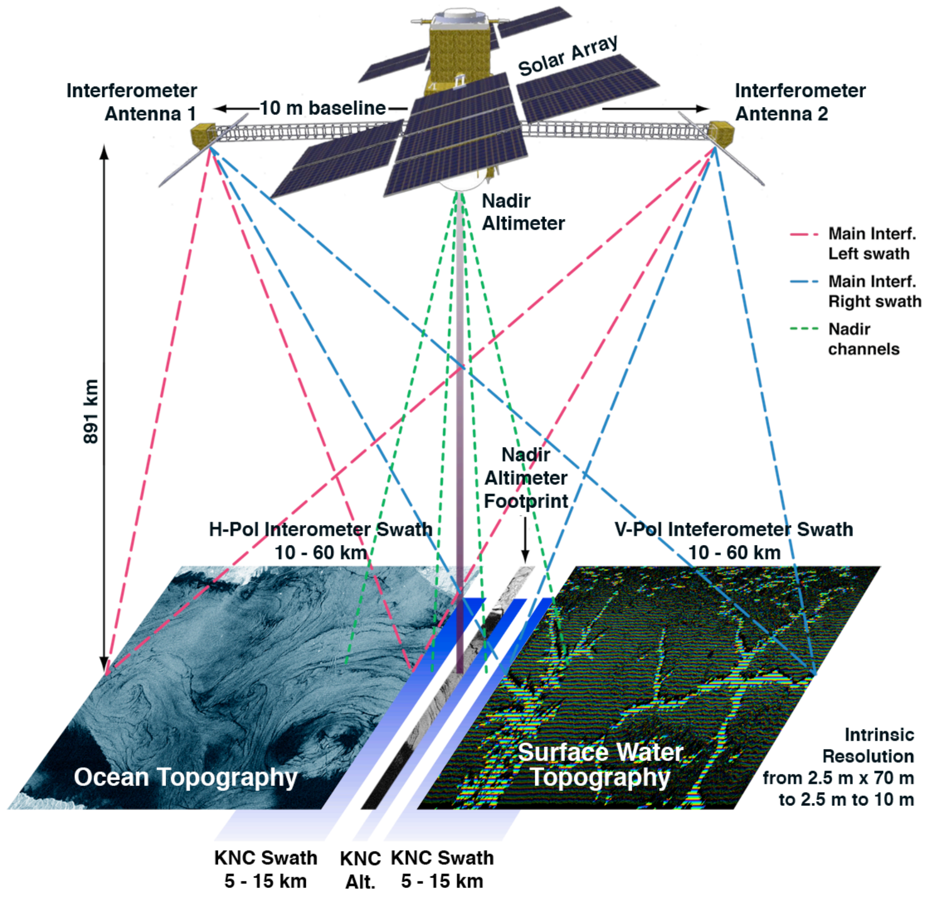
(Garnier et al, 2016)

Assimilation of images

* Images (here, chlorophyll) clearly reveal the structure of the flow;
* How can such data be assimilated into models as images?

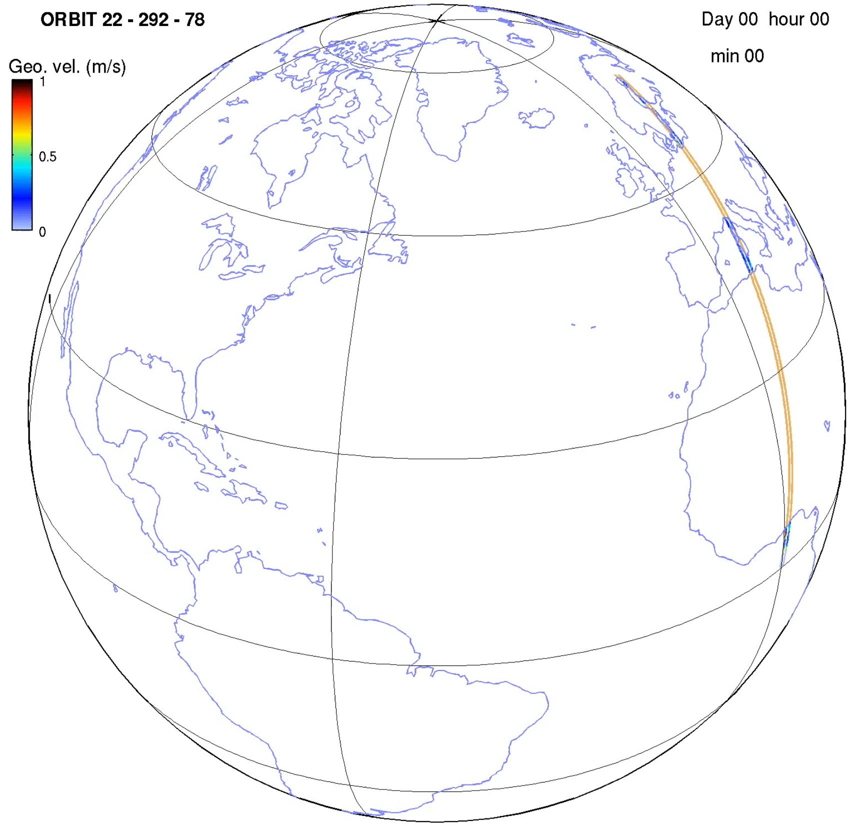
SWOT

- SWOT: Surface Water and Ocean Topography
- Satellite mission to be launched in 2021
- Revolutionary altimetric observation: 120 km-wide swath
- Pixel of 1 km



Future challenges

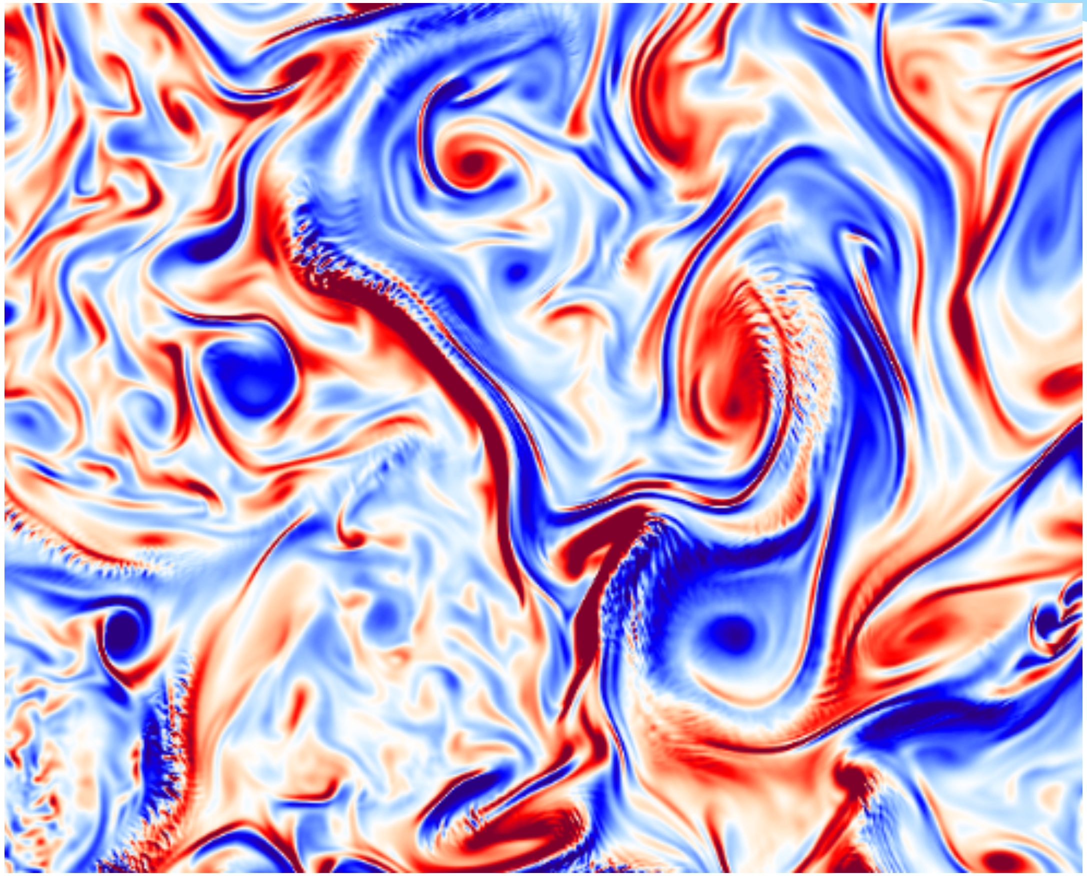
SWOT



SWOT

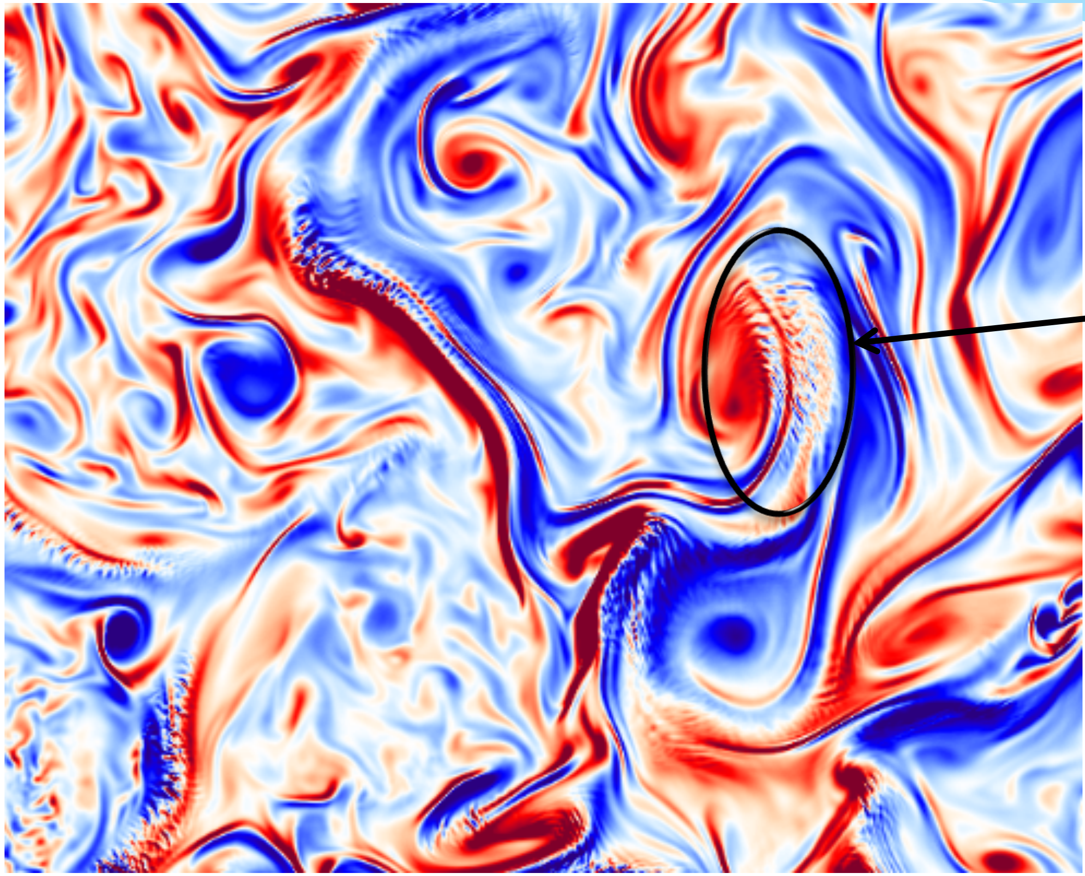
* Challenges:
  * The physical processes that will be observed are not well known;

SWOT



Snapshot of ΔSSH from the 1/60° North Atlantic simulation

SWOT



Physical processes or numerical artefact?

SWOT

- Challenges:
  - The physical processes that will be observed are not well known;
  - The signature of internal tides can be superposed to the balanced dynamics;

SWOT

Phenomenon	Length scale L	Velocity scale U	Time scale T
Atmosphere:			
Sea breeze	5–50 km	1–10 m/s	12 h
Mountain waves	10–100 km	1–20 m/s	Days
Weather patterns	100–5000 km	1–50 m/s	Days to weeks
Prevailing winds	Global	5–50 m/s	Seasons to years
Climatic variations	Global	1–50 m/s	Decades and beyond
Ocean:			
Internal waves	1–20 km	0.05–0.5 m/s	Minutes to hours
Coastal upwelling	1–10 km	0.1–1 m/s	Several days
Large eddies, fronts	10–200 km	0.1–1 m/s	Days to weeks
Major currents	50–500 km	0.5–2 m/s	Weeks to seasons
Large-scale gyres	Basin scale	0.01–0.1 m/s	Decades and beyond

SWOT

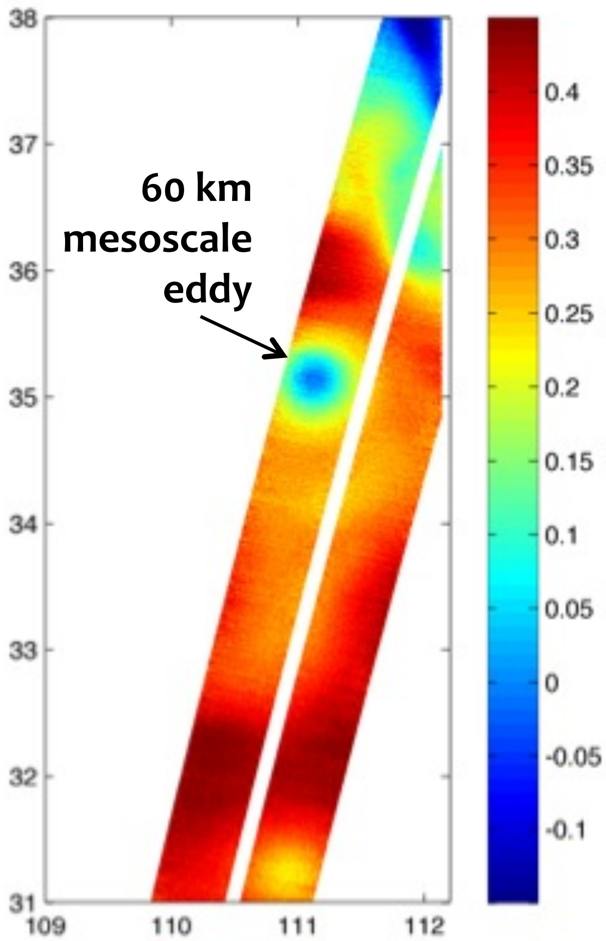
Conventional nadir altimetry

SWOT

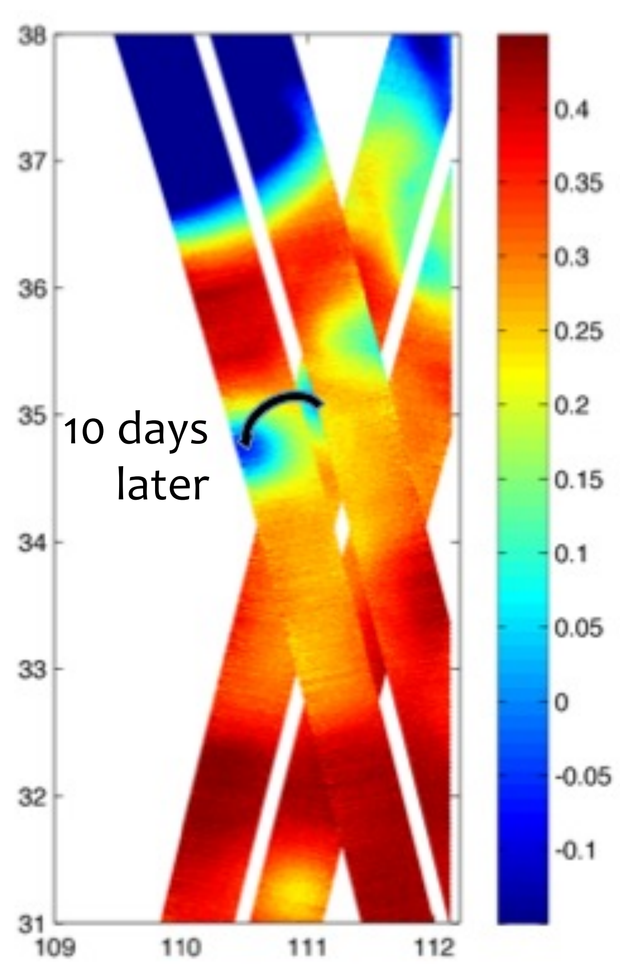
- Challenges:
  - The physical processes that will be observed are not well known;
  - The signature of internal tides cam be superposed to the balanced dynamics;
  - The satellite will provide well separated (in time) snapshots of short-lived structures.

Future challenges

SWOT



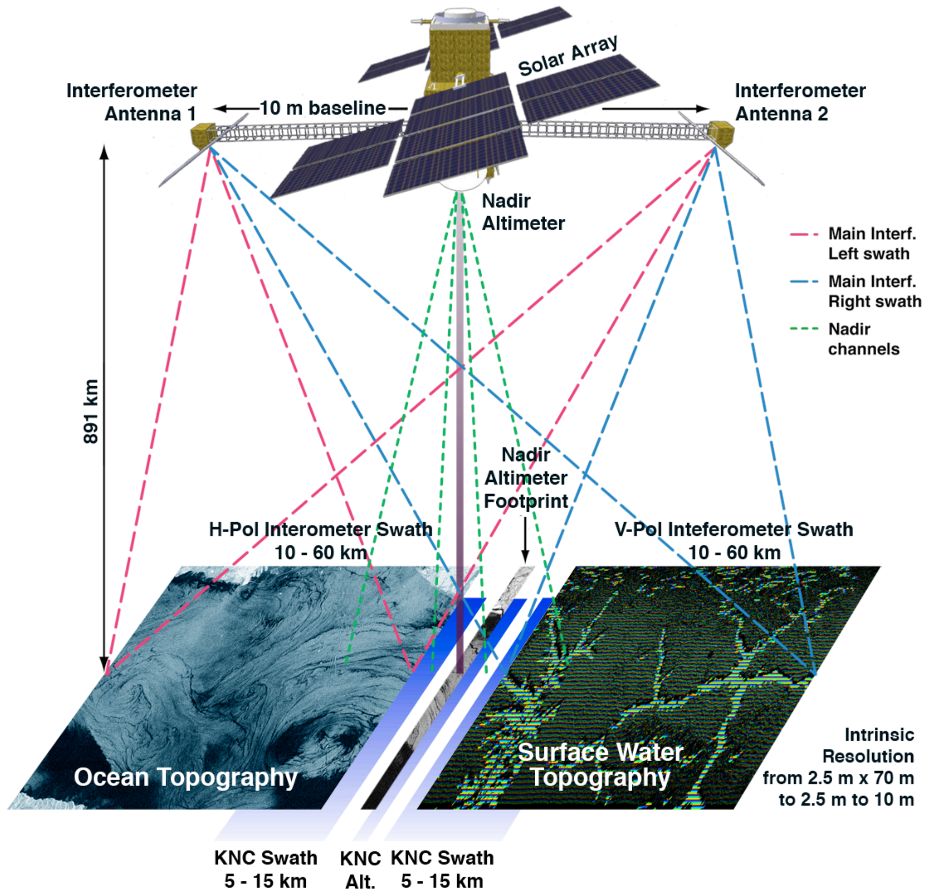
SWOT



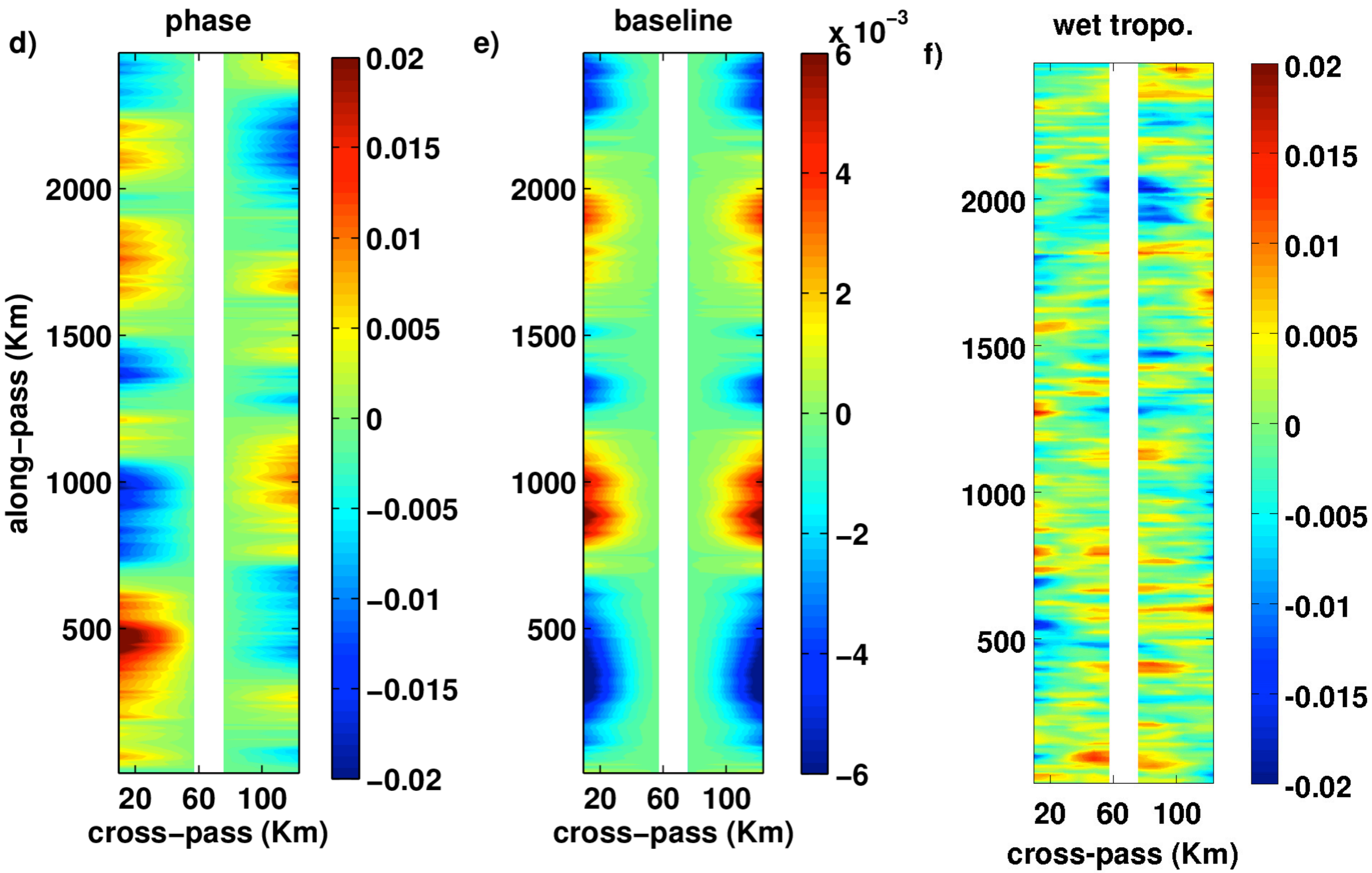
Can we retrieve the SSH evolution between the two satellite revisits?

SWOT

* SWOT observations will be affected by correlated noise (due to roll, tropospheric water vapor, dilation of the baseline, etc)

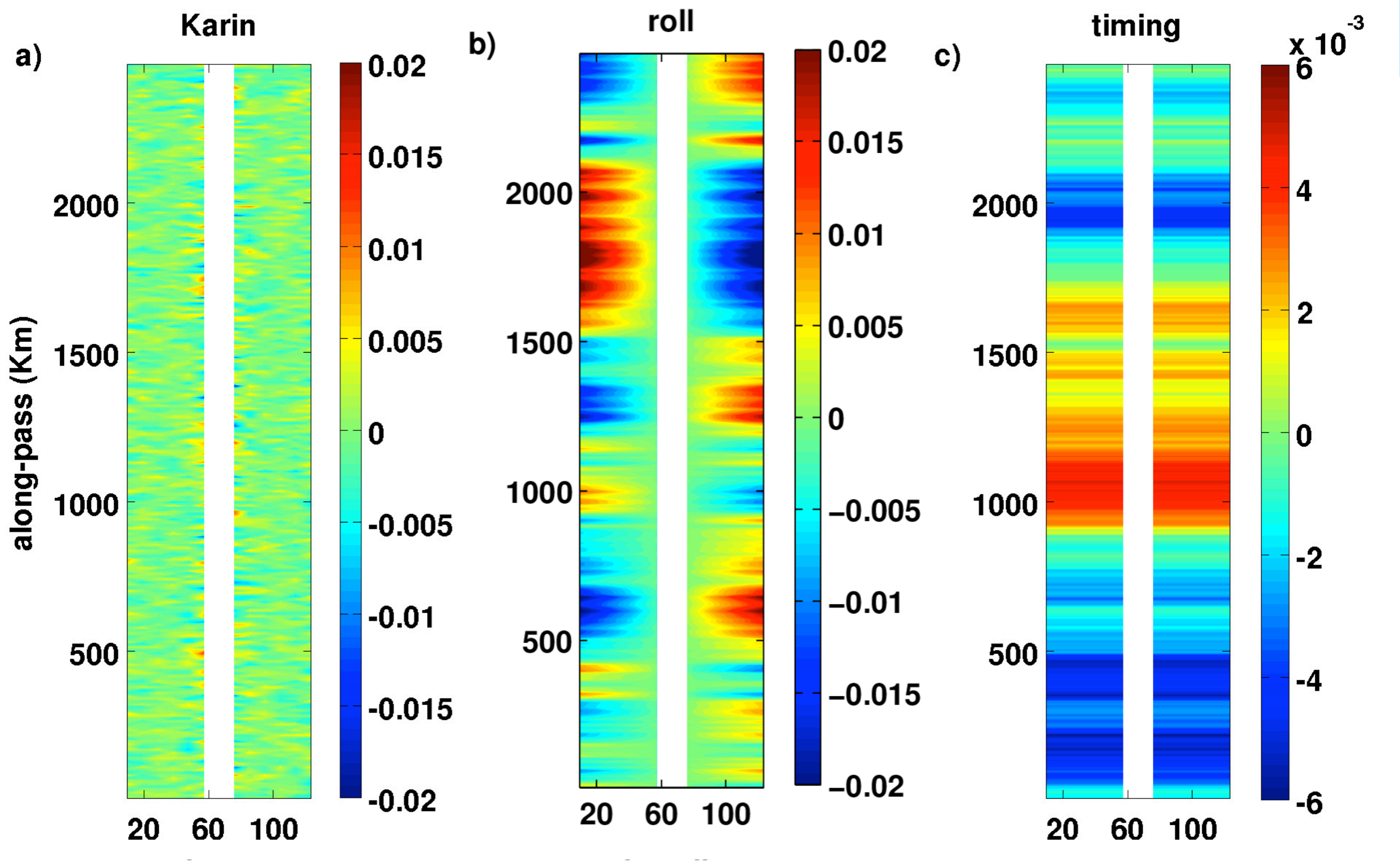


SWOT

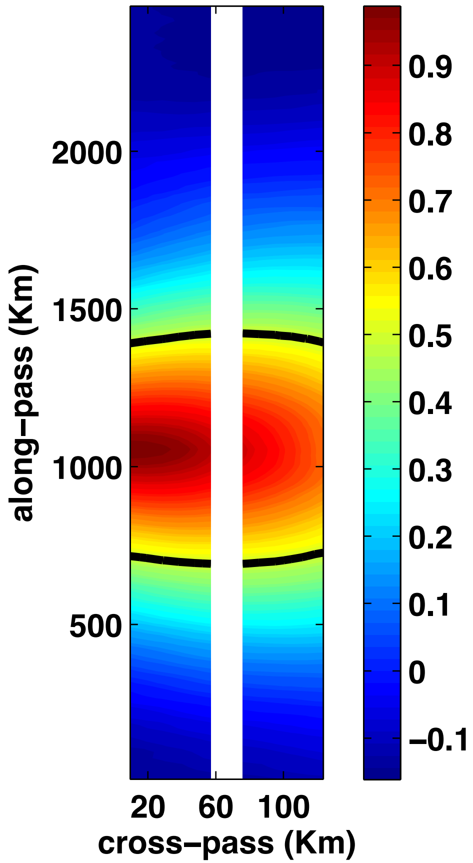

d)
phase
along–pass (Km)
2000
1500
1000
500
20 60 100
cross–pass (Km)
0.02
0.015
0.01
0.005
0
−0.005
−0.01
−0.015
−0.02
e)
baseline
x 10−3
2000
1500
1000
500
20 60 100
cross–pass (Km)
6
4
2
0
−2
−4
−6
f)
wet tropo.
2000
1500
1000
500
20 60 100
cross-pass (Km)
0.02
0.015
0.01
0.005
0
-0.005
-0.01
-0.015
-0.02

Future challenges

SWOT



SWOT



Correlations of errors.

At a 9 km resolution, KaRIn noise is filtered out.

Thank you